Manuscript Information

Journal name:	Expert review of respiratory medicine
NIHMS ID:	NIHMS799428
Manuscript Title:	Assessment of treatment response in tuberculosis
Submitter:	Taylor And Francis Group (gregg.warren@tandf.co.uk)

Manuscript Files

Type	Fig/Table #	Filename	Size	Uploaded
manuscript		1166960.docx	186247	2016-06-30 05:04:37
table	Tables 1-4	Tabs1-4.docx	206509	2016-06-30 05:04:38

Review

Assessment of treatment response in tuberculosis

Neesha Rockwood[1,2]*

Elsa du Bruyn[2]

Thomas Morris[1]

Robert J Wilkinson[1,2,3]

1. Department of Medicine, Imperial College London W2 1PG, United Kingdom
2. Clinical Infectious Diseases Research Initiative, Institute of Infectious Diseases and Molecular Medicine and Department of Medicine, University of Cape Town, Observatory 7925, Republic of South Africa
3. The Francis Crick Institute Mill Hill Laboratory, London NW7 1AA, United Kingdom

* To whom correspondence should be addressed at Department of Medicine, Imperial College London W2 1PG, United Kingdom

e-mail neesha.rockwood@doctors.org.uk tel. +442075943891

fax +27214066084

Keywords

Tuberculosis; Therapeutics; treatment outcome; biomarkers; humans

Abstract

Antibiotic treatment of tuberculosis has a duration of several months. There is significant variability of the host immune response and the pharmacokinetic-pharmacodynamic properties of *Mycobacterium tuberculosis* sub-populations at the site of disease. A limitation of sputum-based measures of treatment response may be sub-optimal detection and monitoring of *Mycobacterium tuberculosis* sub-populations. Potential biomarkers and surrogate endpoints should be benchmarked against hard clinical outcomes (failure/relapse/death) and may need tailoring to specific patient populations. Here, we assess the evidence supporting currently utilized and future potential host and pathogen-based models and biomarkers for monitoring treatment response in active and latent tuberculosis. Biomarkers for monitoring treatment response in extrapulmonary, pediatric and drug resistant tuberculosis are research priorities.

Mycobacterium tuberculosis (MTB) causes a wide spectrum of disease in different patient groups. Up to one third of the world's population are assumed infected with latent tuberculosis infection (LTBI), a proportion of which will progress to active infection. Patients with LTBI are asymptomatic with evidence of persistent immunological sensitization against MTB antigens as evidenced by a positive mantoux tuberculin skin test or positive interferon gamma release assay (IGRA). Reactivation of LTBI can be averted by currently advocated regimens with efficacy ranging from 60-90% [1]. The global human immunodeficiency virus (HIV) pandemic increases both tuberculosis (TB) incidence and TB related deaths, particularly in Africa. In 2014, the World Health Organization (WHO) estimated there were 9.6 million incident TB cases (12% HIV co-infected) and 1.5 million TB deaths (27% HIV co-infected) [2]. In 2014 there was an estimated 480,000 cases resistant to key first line drug rifampicin (RIF) and isoniazid (INH) (multi drug resistant (MDR) TB), of whom approximately 9.7% had additional resistance to fluoroquinolones and aminoglycosides (extensively drug resistant (XDR) TB) [2]. TB treatment is believed to follow a biphasic response to chemotherapy with the majority of bacilli being rapidly killed early on in the bactericidal phase whilst the sterilising phase slowly results in eradication of persisting MTB to reduce risk of relapse. These subpopulations may be anatomically sequestrated and thereby escape immune surveillance and drug penetration. Currently endorsed treatment regimens vary in duration between 6 months in drug susceptible (DS)-TB to 24 months in drug

resistant (DR)-TB. Although treatment success rates in DS TB are generally close to 90% [2], in drug resistant TB they are below 50% [3].

Standardized definitions of treatment response and outcomes enable optimal surveillance and valid comparisons across settings (Table 1).

Priorities in assessment of response to treatment include both identification and prediction of treatment induced adverse events such as drug side effects and paradoxical reactions, temporary and permanent disability secondary to disease, treatment failure and subsequent relapse. Assessment of treatment response on an individualized and programmatic basis in both adults and children is predominantly based on clinical, radiological and bacteriological measures. Clinical assessments include overall performance status, weight gain during treatment and resolution of systemic and organ based symptoms. Radiological assessments include use of plain film radiograph to assess resolution of pathology and to diagnosis of paradoxical reactions. Computed tomography (CT) and/or magnetic resonance imaging are also utilized in certain cases if clinically indicated and within available resource. Bacteriological assessment includes smear and or culture conversion at various intervals depending on whether treating DS- or DR-TB [4]. Blood tests such as C-reactive protein (CRP) can be done at baseline and during treatment to monitor clinical improvement. Monitoring of liver enzymes pre-treatment is indicated in those with risk of premorbid liver disorder and should be monitored, along with all patients symptomatic of drug induced liver injury [5]. In some settings, therapeutic drug monitoring is used [6].

When conducting research, priorities in assessment of treatment response are to identify predictors of long-term outcomes, quantify efficacy of agents in the context of multidrug therapy, and to ascertain non-inferiority of treatment-shortening regimens. In TB, adequate powering of studies, and lengthy follow up times are required to reliably predict long term unfavourable outcomes, which is expensive and logistically challenging. For this reason, biomarkers and correlate endpoints predicting long-term outcomes are desirable [7]. Changes induced by a therapy on a correlate should reliably reflect changes in the long-term treatment outcome [8]. We will review evidence of host and pathogen based biomarkers of long-term treatment response in latent and active TB and outline key considerations for future research.

PATHOGEN BASED MODELS AND BIOMARKERS

Hollow fiber system model

The hollow fiber bioreactor system can be used as an *in vitro* pharmacokinetic-pharmacodynamic (PK-PD) model of TB [9]. It has been used to assess early bactericidal activity and recreate realistic PK patterns of single and multi-drug exposure. It can predict PK/PD exposures associated with development of drug resistance. In tandem with Monte Carlo simulations, it can also predict impact of dose titration and dosing schedules. It can be adapted to reproduce environmental conditions of persistence models to study pharmacodynamics of sterilizing activity [9]. However, it does not model the host immune response, differential spatial distribution in diseased tissue nor the impact of clinical strain variation. Results need to be validated in adequately powered, appropriately controlled clinical studies.

Sputum based assays (summarized in table 2)

1) Early bactericidal activity studies

Early bactericidal activity (EBA) refers to the fall in log10 colony forming units (CFUs) of MTB per millilitre sputum per day. This can also be measured by time to culture positivity (TTP) in liquid media which inversely relates to the number of viable MTB in sputum. Serial measurement of decrease in number of CFU/ increase in time to culture positivity in sputum samples over 2-14 days is used to assess EBA of single drugs or combinations. This chiefly predicts activity against rapidly multiplying extracellular MTB in the wall of cavities. Studies can be performed relatively quickly, results are reproducible and provide data on

bactericidal efficacy of a drug or regimen and its relationship to PK. Hence, appropriate doses and combinations can be taken forward to clinical studies [10]. Unfortunately, EBA studies do not accurately predict activity against non-replicating persisters, sterilization of tuberculosis lesions and risk of relapse. Also, as EBA studies are carried out during the first 2-14 days of treatment, they cannot predict important drug-drug antagonism during sterilization phase. For example, although INH has been shown to contribute most to EBA in combination therapy, studies have shown that INH may have an antagonistic effect on the sterilizing action of PZA, particularly after the first 2 weeks [11]. Limitations in methodology include variability between different centers and intra-individual variation pre-treatment [12]. Patients must be able to expectorate good volumes of sputum. There may be loss of viable CFUs cultured if sputum is decontaminated with sodium hydroxide treatment. There is some evidence to suggest type of culture media and growth factors/inhibitors [13] influence growth of different populations of MTB during chemotherapy [14].

2) 2-month sputum culture conversion and time to culture conversion as intermediate bacteriological endpoints

In a meta-analysis Wallis *et al* found a relationship between 2-month culture status on solid media and relapse rate that was consistent across regions [15]. There was significant correlation between 2-month (Hong Kong) and 3-month (Africa) culture status and combined rate of failure, relapse and death [16]. Johnson *et al* showed that treatment shortening from 6 months to 4 months in patients with both non-cavitatory pulmonary disease and culture negative status

at 2 months was associated with a significantly increased risk of relapse compared with a standard 6 month regimen [17]. These findings were also supported by another retrospective analysis [18]. Despite Phase 2 studies demonstrating higher rates of negative 2-month sputum cultures and earlier time to culture conversion when compared to standard 6-month control regimens [19,20], four phase 3 treatment shortening trials involving fluoroquinolones [21-24] failed to show non-inferiority of treatment shortening. This apparent lack of correlation between sputum conversion and treatment outcome is thought to be due bacterial persistence and incomplete sterilization by 4-month fluoroquinolone regimens. Using meta-regression modeling, Wallis *et al* showed that 2-month sputum conversion and treatment duration could independently predict relapse [25]. This model also predicted that, given the treatment duration of aforementioned trials, there would be unacceptable relapse rates as clinical outcome [25]. The predictive value of 2-month sputum conversion was maintained despite incorporation of data from the fluoroquinolone trials [26].

Time to culture conversion over 8 weeks, in solid (serial CFU counts) or liquid cultures can be modeled to predict bacillary elimination rates (BER) [27]. Sloan *et al* showed BER in the sterilizing phase (after 7 days) was non-linear with high inter-individual variability. A significant negative correlation was seen between sterilization phase BER and treatment failure or relapse [28]. However, no threshold was identified to predict failure or relapse.

Kurbatova *et al* used data from 2 cohort studies to develop models assessing the utility of culture conversion at 2 and 6 months as proxy markers of end of

treatment response in MDR-TB. Culture conversion at 6 months had a significant association with treatment success in MDR TB compared with failure or death (adjusted OR 14.07) with a predicted sensitivity of 92% and specificity of 58%. Culture conversion at 2 months was significantly associated with successful outcome in HIV uninfected patients (adjusted OR 4.12) but had less utility in HIV infected patients and low overall sensitivity [29].

3) Polymerase chain reaction-based methods for quantification of viable mycobacteria in sputum

Polymerase chain reaction (PCR) based methods with real time detection of MTB deoxyribonucleic acid (DNA) have had a significant impact on programmatic diagnosis of MTB and baseline drug resistance. The utility of using tests such as GeneXpert® MTB/RIF (Cepheid Sunnyvale, CA, USA) and GenoType MTBDR*plus*/MTBDR*sl* (Hain Lifescience Nehren, Germany) to ascertain sputum sterilization at any point during treatment is limited by DNA amplification from dead bacteria in clinical samples [30]. There is ongoing research to improve specificity of DNA amplification from live bacteria through treatment of sputa with agents such as propidium monazide, a DNA-binding dye that penetrates through damaged cell walls and inhibits PCR amplification via DNA modification [31]. An alternative approach includes quantification of abundant ribonucleic acid (RNA) species extracted from sputum as surrogate markers of bacterial clearance [32,33]. Honeyborne *et al* showed a high correlation between 16S ribosomal RNA (rRNA) with TTP in liquid media and CFUs on solid media over the first 14 days of treatment [34]. Although the half-life of messenger RNA (mRNA) and

rRNA is significantly shorter than DNA in sputum, in the context of ongoing chemotherapy, it is unclear what proportion of m- and rRNA in expectorated sputum originates from dying bacteria. Although results suggested by biexponential modeling are promising, further studies are required to verify if measurements of 16s rRNA earlier on in treatment can predict failure and relapse [35]. Clinical studies are required to ascertain if detection of rRNA and/or mRNA in the sputum of clinically asymptomatic culture negative patients at the latter stages of treatment i.e. potentially viable but non-culturable MTB, correlates with relapse. Clearance of 85B mRNA by Day 2 was also associated with early culture conversion by 1 month [32].

4) Staining of sputum (viability and lipid bodies)

Assessment of MTB viability early during treatment may expedite detection of true (culture proven) treatment failure [36]. This may be appropriate in settings with limited laboratory capacity, and infrastructure to carry out culture based techniques and drug susceptibility testing. Fluorescein diacetate, used to stain clinical samples, is hydrolyzed by acetylesterase by live metabolically active MTB and resulting fluorescent bacilli quantified using a light emitting diode fluorescence microscope. Datta *et al.* showed quantitative viability microscopy accurately predicted the concentration of culturable MTB in sputum and was able to differentiate DS and DR TB patients within 9 days of treatment [37]. This technique is not applicable in smear negative patients. Further potential utility of this viability staining technique includes monitoring efficacy of treatment response to both novel and MDR TB treatment regimens.

MTB stores and utilizes lipids as an energy source. *In vitro* models of dormancy have shown accumulation of host triacylglycerol as lipid bodies under control of triacyl glycerol synthase 1 [38]. These lipid bodies can be detected in sputum via staining with auramine-labelled Nile-Red stain both pre- and during treatment. Change in detectable lipid bodies during the first 2 weeks of treatment, is hypothesized to reflect an increase in viable but non-culturable persisters, and significantly varies with different drugs [39]. In a preliminary study, Sloan *et al.* showed that at week 3-4 of treatment, there was 21% increased odds of unfavourable outcome (failure/relapse) for each percentage rise in % lipid body positive acid-fast bacilli (AFB). Baseline counts of % lipid body positive AFB did not predict treatment response [28].

5) Whole blood bactericidal assay

Growth of clinical MTB isolates in *ex vivo* whole blood culture taken at selected time points during therapy is a potential biomarker of sterilizing activity, encompassing both strain variation and host effector immune mechanisms. In one study, whole blood bactericidal activity correlated with rate of fall in sputum CFUs over the first 4 weeks of antituberculosis treatment (ATT) and inferior bactericidal activity was seen in whole blood culture of patients with delayed sputum sterilization [40]. Whole blood culture of H37Rv at different doses and in combination with novel regimens for DR-TB has been used to rapidly assess bactericidal potential and drug-drug antagonism/synergy [41].

6) Mycobacterial products (non-nucleic acid)

Wallis *et al.* showed that levels of Antigen 85 (Ag 85) in sputum, as quantified by ELISA, with levels > 60pg/ml at 2 weeks were predictive of liquid cultures becoming positive within 20 days beyond day 90 of therapy. This microbiological feature identified both of the patients that failed treatment in this study. Conversely, Ag 85 < 60pg/mL at 2 weeks was predictive of rapid cure in 26 of 26 cases [42]. The accuracy and reliability of Ag 85 measurement by this method and its ability to reflect initial bactericidal activity was confirmed in a further study [43].

Non-sputum based mycobacterial products

The other antigen-body fluid combination that has been studied during treatment is lipoarabinomannan (LAM) in urine. Wood *et al.* measured urine LAM by ELISA in 200 adults with active TB pre-treatment, daily during week 1, then weekly at weeks 2, 8, 16 and 24 [44]. LAM positivity was almost entirely restricted to individuals co-infected with HIV-1 co-infection and with low CD4+ counts (especially < 50 cells/ul). Average LAM levels throughout the LAM-positive cohort remained the same for the first 2 weeks of treatment but then dropped by around 1 $\log_{10}$ by 8 weeks, disappearing altogether by 24 weeks. Use of this marker for clinical trials would require further validation, however, as levels of LAM in urine may largely reflect bacterial load in the kidney rather than systemically or in the lung. Further verification has been provided to some extent by a pilot study by Drain *et al.*, in which 29 urine LAM-positive adult patients in Durban, South Africa, had repeat assays performed at 2 and 6 months [45]. LAM levels decreased significantly from 0 to 2 months and again through 6 months, but the

sample size was not adequate to correlate LAM levels with clinical improvement. LAM was measured using the Alere Determine^{TM} TB LAM Ag test that is semi-quantitative, offering a choice between one of 4 categories of positivity. A LAM grade of 2 or more at 2 months was associated with a hazard risk (HR) of death from any cause of 5.58; and any positivity at 6 months with a HR of 42.1, however, although cause of death was not known in every case.

In latent TB, Young *et al.* found 6 mycobacterial proteins that were present in urine of HIV uninfected patients which suggests that products other than LAM may be useful for measuring bacterial load over a wide dynamic range, but these require verification in longitudinal studies [46].

HOST BASED BIOMARKERS

Lung function testing

Most active pulmonary TB patients exhibit a combined obstructive/restrictive pattern of lung function impairment, however the restrictive impairment pattern seems to be more responsive to successful ATT with the residual loss in lung function being predominantly obstructive in nature [47,48]. A South African study evaluating the influence of ATT on spirometry, radiographic score and inflammatory markers in hospitalised pulmonary TB patients found that impairment in lung function did improve with ATT in most (54%), but that a substantial proportion of pulmonaryTB patients developed residual impairment (26%) or worsened outcome (20%) in lung function in comparison to pulmonary function at treatment initiation [48]. Interestingly participants who were HIV-1 infected failed to demonstrate significant improvement in lung function at treatment completion. This was independent of smoking status and adherence to treatment. Furthermore the extent of loss in lung function, correlated with number of TB episodes [49]. It can therefore be inferred that resolution of a restrictive lung function pattern may be the most useful spirometric correlate of treatment response in the acute setting, whereas the degree of airflow obstruction at treatment completion and thereafter has value in determining the degree of residual impairment that may lead to chronic airflow obstruction and disability.

Radiological markers (summarized in table 3)

1) Chest radiography

Chest radiography (CXR) is widely utilized in patients with TB symptoms although it lacks specificity as diagnostic test. Problematic factors in the application of chest radiographs as both a diagnostic and monitoring modality include inter-observer variability, atypical radiographic appearance of TB in the setting of HIV-1 co-infection and lack of a universally reading tool/scoring system for either. The Timika score that comprises a simple numerical score that grades radiographic severity in the setting of smear positive TB, has found application particularly in the clinical trials setting as predictive of 2-month sputum conversion status. The score comprises the proportion of visible pulmonary involvement (%) with the number 40 added should one or more cavity be present [50]. Although it is well established that there is an inverse association between cavity number and volume on chest radiograph and TTP in liquid culture media [51], it was still found that the Timika score out-performed its individual components in predicting 2 month sputum status. At a cut-off of 71 the score could predict a positive sputum smear at 2 months with a sensitivity of 80% (95% CI 61.4 to 92.3) and a specificity of 67.7% (95% CI 57.3 to 77.1). There was positive correlation between the score and smear grade at diagnosis ($p<0.001$) and quality of life assessment by St George's Respiratory Questionnaire (higher scores indicative of greater impairment) and a negative correlation with body mass index, forced expiratory volume in one second or FEV1 (% of predicted), haemoglobin, and exercise tolerance as measured by the 6-minute walk test ($p<0.02$ for all correlates other than smear grade). The score had initially been developed and validated in an Indonesian population that was mostly HIV-1

uninfected, however in an unrelated cohort from Cameroon with 30% HIV-1 co-infection the score also predicted sputum non-conversion at two months with similar specificity and sensitivity, albeit by using a lower cut-off value [52]. Although there was poor inter-reader agreement in the initial study by Ralph *et al.* (prevalence and bias adjusted kappa: 0.37 for cavitation, 0.31 for patchy consolidation and 0.7 for confluent consolidation), a recent publication from South Africa showed greater agreement (kappa value for cavitation: 0.66, overall inter-reader correlation: $r=0.86$, $p<0.001$) in using the Timika score [53]. The latter study however found an optimal score cut-off of 61.3, which could only predict 2 month sputum smear status with a sensitivity of 74.1% (95%CI 65.0 to 81.9) and a specificity of 57.7% (95%CI 51.7 to 63.6). Utility in predicting long term outcomes of failure and relapse by month 30 was limited by suboptimal sensitivity and specificity of 67% and 57% respectively. The persistance of one or more cavity on chest radiograph after treatment completion may have prognostic value. Indeed it was found that patients with a persistent cavity on their end of treatment CXR had more than twice the risk for TB relapse when compared to those with resolved cavities [54].

2) Ultrasound

Particularly in the context of asymptomatic disease, ultrasound (US) represents an imaging modality with reliable detection capabilities that may be utilized by the clinician and researcher alike. The commonest radiographic finding in pulmonary TB in children is mediastinal lymphadenopathy, which may be readily detected by US. US detection of mediastinal lymphadenopathy was comparable to CT and

had superior sensitivity to chest x-ray [55]. Bosch-Marcet *et al.* performed a retrospective study of 21 children who had been treated for active TB on the basis of a positive tuberculin skin test and mediastinal lymphadenopathy following suprasternal or left parasternal ultrasound scanning [56]. Only 11 (52%) were symptomatic. US was repeated at 3 months, and in 17 (81%) cases, the extent of lymphadenopathy decreased by at least one category (*5 ad hoc* categories had been established by the authors purely for the purposes of this study) with 2 participants responding later. The importance of these findings is slightly blunted by the fact that the cut-offs used to define a response may have been fitted to the results. Also, US is non-invasive with low running costs. Further studies, including both HIV infected and uninfected children and earlier time points, are warranted.

Sharma *et al.* employed US as part of a monitoring package to assess the efficacy of thrice-weekly ATT for 351 patients with definitive or probable tuberculous pleural effusions in India. All cases were HIV uninfected with uncomplicated small (<1.5 l) unilateral pleural effusions [57]. Clinical and sonographic assessments were performed at 0, 2, 4 and 6 months and clinical follow-up was also performed post-treatment at 9, 12, 18 and 24 months. Of the 308 patients that completed follow-up to 6 months, 89% had "complete resolution" as defined by clinical and sonographic recovery. A negative US at 6 months was therefore specific for cure but not 100% sensitive, as of the 26 patients that still had residual fluid at 6 months, only 2 relapsed. Ultrasound also allowed differentiation of residual fluid from pleural thickening. The majority of

cases did not require drainage and so comparison with pre-treatment effusion volume is relevant. Follow up data were not shown in this study. Studies informing the utility of US in monitoring TB treatment response in HIV co-infection are scarce.

With regard to peritoneal TB, US may be useful for follow-up, having confirmed the diagnosis by more specific means, but there are currently relatively few studies supporting this [58,59]. Further studies at earlier time points, with larger numbers and in different age groups and degrees of immunosuppression are warranted.

3) Combined ^{18}Fluorodeoxyglucose positron emission- and computerized axial- tomography (PET/CT)

PET/CT is an established method for monitoring response to cancer treatment and has recently been incorporated into clinical studies of TB. Chen *et al.* performed PET/CT imaging (high resolution CT) as a sub-study of 35 adults with MDR-TB who had been enrolled into a trial of adjunctive metronidazole vs. placebo in South Korea [60]. The results showed that, in comparison to sputum culture status or CT alone at 2 months, PET/CT performed better with a sensitivity of 0.96 and specificity of 0.75 to predict durable cure. This was on the basis of change from baseline activity/score. This study provided proof of concept that PET/CT may provide a valuable early outcome measure for use in clinical trials. PET and CT appeared highly complementary in that whilst cavities (highly predictive of outcome) are cold on PET, other lesions such as nodules or

consolidations are better quantified by metabolic activity (C.E. Barry III, private communication).

Coleman *et al.* studied PET/CT to monitor of XDR-TB infected macaques at 0, 1 and 2 months after commencement of treatment with linezolid (LZD) monotherapy [61]. The treatment group showed a significant reduction in disease activity by PET as early as 1 month. The authors also analysed PET/CT results from a previously published human study in which LZD was added (successfully) to the regimen of 41 patients with XDR-TB who were failing treatment. The majority of patients who received repeat imaging within 3 months post-LZD had decreases in their PET activity scores of magnitude similar to that seen in macaques. This provided validation of the macaque model for assessing LZD potency. In addition, 1 participant who had an initial response by PET at 1 month had a subsequent increase in FDG uptake at 6 months, correlating with the emergence of LZD resistance. PET/CT therefore appears to be a promising technique for predicting outcome of TB treatment, possibly as early as 1 month. Further studies in humans employing PET/CT at 1 month and 2 weeks are in progress. A limitation to repeated use of this imaging modality is the high level of radiation exposure when combined with CT. Also, the assay measures total metabolic activity at the disease site thought mainly to reflect host neutrophil response.

Ghessani *et al.* performed PET/CT before and after treatment of "latent TB" in 5 adult close contacts with positive IGRAs [62]. 3 of the 4 participants who had detectable FDG uptake in hilar lymph nodes showed a decrease or resolution of

uptake post-treatment, whilst the 5th only showed calcification on CT initially. Whether or not these changes occurred as a result of drug therapy or merely reflect the natural history of LTBI, this small study indicates that PET/CT does have the resolution to enable this area to be explored in greater detail. PET/CT may have potential to become a gold standard for monitoring response to LTBI treatment in clinical trials and for assessing the performance of other, simpler methods of judging position of individual patients on the spectrum of LTBI.

Immune markers (summarized in table 4)

1) Acute phase proteins

The acute-phase response can be triggered by inflammation, infection or tissue injury and is characterised by cytokine induced release of predominantly hepatic synthesised proteins into the circulation. Although not specific to TB, a number of the acute-phase proteins have been proposed as useful in monitoring the decline in inflammation typically observed in response to antimicrobials. Most notably, CRP concentration rapidly decreases within the first month of ATT in tandem with the reduction in sputum bacillary load observed in treatment responsive patients [63] and return to normal of initially elevated CRP concentrations may correlate with therapeutic response [64]. Conversely, pre-treatment concentrations above 20mg/l which persisted during treatment was associated with adverse treatment outcomes in one small study [65]. Procalcitonin (PCT) levels are rarely significantly elevated in tuberculosis, so much so that it is often used to differentiate acutely between pulmonary TB and pneumonia, where the latter often exhibit elevated levels of PCT. The exception being in cases of severe and

disseminated TB, where PCT levels equal or in excess of 0.5 ng/ml is indicative of poor prognosis and increased mortality risk [66].

2) Cytokines and T lymphocyte subsets

Despite the wide variability and often contrasting findings observed in the review by Clifford *et al.* [67], it was found that the majority of studies evaluating both TB-antigen stimulated and unstimulated cytokine responses found decreasing concentrations of tumour necrosis factor-α (TNFα) over the course of successful ATT. There are contrasting results in the literature regarding the change in IL-10 production during treatment [67]. Sahiratmadja *et al* measured IL-10 and IFNγ response to *ex vivo* stimulation of peripheral blood mononuclear cells with MTB and found evidence of an increasing IFNγ/IL-10 during chemotherapy [68]. Mihret *et al*, also showed an increase in IFNγ/IL-10 in plasma from HIV-1 uninfected (but not HIV-1 co-infected) participants during chemotherapy [69]. Initial data by Harari *et al*, showed that by the end of successful treatment, individuals exhibited a shift from MTB-specific CD4 T cells of single positive TNFα expressing phenotype towards a polyfunctional MTB-specific T cell profile (expressing IFN-γ^+ , TNFα^+ and interleukin-2^+ (IL-2^+)) [70]. It has been suggested that mycobacterial load influences phenotypic expression of MTB-specific CD4 T cells, as single positive TNFα producing T cells are associated with high mycobacterial load. Successful ATT decreases mycobacterial load and restores peripheral T cell proliferation capacity and some studies have shown that polyfunctional T cells and MTB-specific cells T cells with single expression of IL-2 predominate [71]. Some subsequent studies have failed to conclusively

demonstrate that polyfunctional MTB-specific T cells are definitively associated with successfully treated TB, whilst others assert that IL-2^{+} single producing and/or IL-2^{+} and IFN-γ^{+} double producing MTB-specific T cell frequency is associated with successful treatment [72-74]. Multiple studies have shown a decline during therapy in the proportion of unstimulated and MTB antigen-stimulated regulatory T cell subsets ($CD4^{+}CD25^{high}CD127^{low}$, $CD4^{+}CD25^{+}FoxP3^{+}$, $CD4^{+}CD25^{high}CD147^{++}$ and $CD4^{+}CD25^{high}CD127^{low}CD161^{+}$ T_{regs}) in both pulmonary and extrapulmonary TB [75]. This is a potential biomarker in sputum-sparse and smear negative individuals [75]. However, there is variation in T_{reg} responses during chemotherapy in different studies. Differential human leukocyte antigen (HLA) expression in different populations may explain this variation [75]. A recent study also showed that MTB-specific T cell expression of activation status as measured by HLA-DR and CD38, along with the intracellular proliferation marker Ki-67 correlates with both bacillary burden and subsequent sputum conversion [76], but these findings need to be validated in larger prospective studies.

Urokinase-type plasminogen activator receptor is expressed by immune cells in response to bacterial phospholipases and pro-inflammatory cytokines. Rabna *et al.* showed that high plasma levels of soluble urokinase plasminogen activator receptor (suPAR) at inclusion, or at any point during treatment was associated with increased mortality. An increase in suPAR after 1 month, compared with diagnosis, was associated with a Mortality Rate Ratio (MRR) of 4.5 (95%CI: 1.45–14.1) during the remaining 7 month treatment period. However, this study

did not identify a threshold to define ‘high suPAR’ levels but specified that levels below 3.5ng/ml were associated with low mortality [77].

3) Tissue destruction and remodeling markers

There is interest in the matrix metalloproteinases (MMPs) in the setting of TB biomarker research. The MMPs constitute a group of endopeptidases secreted by a variety of host cells in response to TB-related MMP upregulation and in so doing affects enzymatic degradation of extra-cellular matrix proteins which may play a crucial role in cavity formation [78]. Numerous MMPs, most notably MMP-1 and MMP-3, have been found to be increased in patients with active pulmonary TB [79], independently associated with higher TB severity scores [80] and undergo rapid decline during treatment [78]. MMP-1, -3 and -8 concentrations in sputum has also been found to decline in response to successful ATT [81]. Delayed sputum culture conversion was associated with increased MMP-1 levels [82]. Elevated MMP-9 levels have been associated with TB meningitis, with treatment related decline in concentrations being apparent in one small study [83]. Thus, MMPs represent an attractive possibility as TB biomarkers, given the potential of recognition of phenotypically distinct clinical presentations of TB perhaps in conjunction with other TB biomarkers such as heme-oxygenase 1 [84].

Transcriptomic profiling

TB is characterized by an interferon-inducible neutrophil-driven peripheral blood transcriptional signature [85]. Several studies have documented this signature relates to disease extent and shows reversion during successful therapy towards

that of latently infected control participants, with the most profound changes occurring within the first weeks of treatment [85], [86]. Cliff *et al.* showed down-regulation of complement and interferon- related genes involved in the inflammatory response after 1 week. There was a slower up-regulation of lymphocyte components, including B- and T-lymphocyte- related genes between weeks 4 and 26 of ATT [87]. Participants experiencing subsequent relapse showed significant up-regulation of cytotoxic cell-mediated killing in response to MTB *in vitro* both at baseline and up to 4 weeks of ATT [88]. Thus, there is the potential of transcriptomic profiling to aide both research on novel tuberculosis treatments and to monitor therapy. However studies of greater power and with adverse outcomes are required to benchmark this promising technique better.

Interferon-gamma release assays (IGRA)

The antigen specific release of interferon-gamma in response to MTB antigens *in vitro* has become an accepted alternative to the tuberculin skin test when determining immune sensitization. It is also well recognised that change in the response to tuberculosis antigens *in vitro* occurs during ATT [89]. However a recent systematic review on the use of IGRA for treatment monitoring concluded that whilst the response tended to fall during treatment, there was a large degree of individual variation and thus IGRA were not felt useful in this respect [90]. A number of experimental studies have also documented clear changes in antigen specific interferon-gamma release during the course of treatment for latent tuberculosis [91], [92], [93]. However the absence of a ‘gold standard’ for the

diagnosis of latent tuberculosis confounds interpretation and large-scale studies with clinical endpoints have not been performed.

EXPERT COMMENTARY

Nahid *et al.* suggested recommendations for a TB specimen and data repository associated with clinical studies that have been adopted in part by some agencies funding tuberculosis cohorts. These include formation of multi-site consortia with appropriate field site and laboratory equipment, infrastructure and storage facilities. Samples should be taken pre-treatment and longitudinally during an adequate length of follow up [8]. Suggested samples for biobanking including serum, plasma, DNA, RNA, sputum, urine, peripheral blood mononuclear cells (PBMC) and, where appropriate, 'site of disease' samples such as pleural, pericardial, bronchoalveolar and cerebrospinal fluid. Sequential culture isolates should be stored for strain typing and DNA fingerprinting. Detailed clinical metadata should be accrued concurrently, along with adherence assessment. Cohorts from diverse epidemiological settings and representing the full spectrum of paediatric TB should be followed up longitudinally in a standardized manner to create biorepositories enabling development of biomarkers for treatment monitoring in paediatric TB [94].

To date, there is a deficit of studies addressing potential biomarkers to predict treatment failure and relapse in DR-TB. This is a research priority to be addressed in both prospective cohorts and randomised controlled trials assessing efficacy of new DR-TB regimens. Such biomarkers/surrogate endpoints will hopefully expedite regulatory approval of new drugs [95].

An optimized composite disease severity score could be developed to screen individuals at high risk of unsuccessful outcome at baseline or relatively early during treatment. These individuals can then be prioritized for intensive treatment monitoring and potential treatment intensification. This should be ideally based upon multiple minimally invasive modalities that are cost effective, of high sensitivity and have a relatively rapid turn around time. In light of significant inter-reader variability, computer-assisted algorithms of radiographic images may hold promise and deserve further attention.

There is considerable heterogeneity observed in the immune response to tuberculosis with a myriad of host and microbial related factors playing a role in inter-individual and geographic variation in the measured immune response to MTB [96]. Whilst biomarkers should ideally be generalizable across ethnically diverse populations, there is evidence that they need to be tailor-made to specific sub-groups. At this time there is little evidence for any single or combination of cytokines, chemokines or TB-specific cells that can serve as reliable surrogate marker of TB treatment response.

In pulmonary TB, there is the possible limitation of - detection and monitoring of viable MTB sub-populations including minority drug resistant variants and non-replicating persisters. Expectorated sputum is a stochastic sampling of the disease site of varying inoculum size. In the laboratory, variability in decontamination practices, use of media and supplements and MIC breakpoints can significantly affect surrogate endpoints. Where possible, universal standard operating procedures should be prepared when planning large multi-site studies.

The PanBiome Study by the PanACEA and PreDiCT-TB consortia (http://www.predict-tb.eu/panbiome-project) will report novel molecular and culture approaches as biomarkers of treatment response, as compared with traditional solid and liquid culture techniques. The availability of high coverage genome sequencing directly from clinical samples may assist optimal characterization of minority bacterial subpopulations. Novel methods of characterizing the time-kill effect of chemotherapy on bacterial sub-populations over the dosing interval could be developed through pharmacodynamic modeling.

Transcriptomic profiling could contribute significantly to the research evaluation of novel or repurposed antimicrobial and host-directed drugs. However the methods are not standardised and studies that benchmark against sufficient hard clinical outcomes are necessary. Ongoing work of promise revolves around the definition of a transcriptomic signature of risk of progression in latently infected persons.

Presently available commercial interferon-gamma release assays have little role in monitoring the treatment of active TB. Whether experimental assays incorporating new antigens and cytokine combinations could contribute to monitoring the treatment of active or LTBI is unknown. The latter would be very helpful as there is no way to infer the likely efficacy of a treatment for LTBI other than by a very large trial with the clinical endpoint of active TB. Ongoing studies of PET/CT in LTBI also have the potential to contribute to research evaluation of new treatment regimes.

FIVE YEAR VIEW

Ultimately, establishing reliable evidence-based measures of monitoring treatment of TB will require a deeper understanding of pathogenesis, including both host and mycobacterial factors.

Novel technologies and integrated use of transcriptomic/proteomic/metabolomic approaches may enhance treatment monitoring, particularly of special treatment groups such as paediatric and extrapulmonary TB. Developing mass spectrometry-based technologies as well as novel biosensors and “antigen capture” techniques such as aptamers are also likely to be investigated.

A preliminary study by Nahid *et al* used a multiplexed aptamer-based proteomic technology to define a non-culture based 5-marker signature predictive of 2-month culture status. This approach shows promise and should be validated in larger cohorts [97]. The diversity of the MTB secretome and the distinctive MTB-associated cell wall lipids provide an attractive option for future sputum antigen based tests. A limitation of these respiratory assays is that they are restricted to populations of bacilli that are able to access the airways. Rapid clearance of bacteria from the sputum is highly desirable and essential, but achievement of this status does not guarantee avoidance of relapse.

An objective of future research is to enhance understanding of post-transcriptional and post-translational regulation of cytokines/chemokines and their role in host-pathogen immune response.

Future research objectives also include advancing knowledge of how lung and sputum microbiome affects MTB response during chemotherapy. Genetic [98]

and epigenetic [99] mechanisms are likely to contribute to treatment response through numerous pathways such as plasticity of the immune response and pharmacogenomic variability. As demonstrated by Prideaux *et al*, differential spatial distribution and kinetics of accumulation in diseased tissue can help predict sterilizing activity of existing and new drugs [100]. Enhanced understanding of the tissue micro-environment will guide rationale for individualized dosing and duration of specific drugs based upon the kinetics of drug action. A multi-systems approach has been used by Pienaar *et al.* to integrate spatio-temporal dynamics of granuloma formation and immune function, PK in plasma and tissue and time to sterilization of intracellular, extracellular replicating and extracellular non-replication MTB populations [101].

KEY ISSUES

- Biomarker research may be enhanced by the standardized collection and preservation of samples longitudinally during treatment. Potential biomarkers should be benchmarked against hard clinical outcomes.
- Some biomarkers may need to be tailored to specific population sub-groups. Biomarkers for monitoring treatment response in paediatric TB and drug resistant TB are a research priority.
- A limitation of sputum-based measures of treatment response may be suboptimal detection and monitoring of viable MTB sub-populations including minority drug resistant variants and non-replicating persisters.
- PET/CT may provide a valuable early outcome measure for use in clinical trials of drug sensitive, drug resistant and latent TB. Ultrasound scanning may be useful in monitoring mediastinal lymphadenophathy in children and treatment response in extrapulmonary TB. A validated chest x-ray score of disease severity at baseline show promise for predicting interim treatment outcomes.
- Currently there is no single or combination of cytokines, chemokines or TB-specific cells which is a reliable surrogate marker of treatment outcomes.
- Presently available interferon-gamma release assays have little role in monitoring the treatment of active TB.
- The blood transcriptomic signature in active TB may predict host response to treatment and correlates with disease severity. Standardization of

methodology is important.

- Interplay between the spatio-temporal distribution of drugs at the disease site, bacillary sub-populations of varying levels of drug susceptiblity and interactions with the host immune microenvironment may be studied via a multi-systems approach.

Declaration of Interest

RJ Wilkinson acknowledges the following funding sources: Wellcome Trust (UK): 084323, 104803; MRC (UK): U1175.02.002.00014; FP7-HEALTH-F3-2013-305578; MRC-SHIP-02-2013, National Institutes of Health 1U01AI115940-01 and National Research Foundation of South Africa (reference 96841). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

References:

•Of interest

••Of considerable interest

1. Lobue P, Menzies D. Treatment of latent tuberculosis infection: An update. *Respirology*, 15(4), 603-622 (2010).
2. World Health Organization. Global Tuberculosis Report 2015. (2015). http://www.who.int/tb/publications/global_report/gtbr2015_executive_summary.pdf?ua=1. Last accessed 09/01/2016.
3. Jacobson KR, Tierney DB, Jeon CY, Mitnick CD, Murray MB. Treatment outcomes among patients with extensively drug-resistant tuberculosis: systematic review and meta-analysis. *Clin. Infect. Dis.*, 51(1), 6-14 (2010).
4. World Health Organization. Definitions and reporting framework for tuberculosis– 2013 revision. http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf?ua=1. Last accessed 09/01/2016.
5. Saukkonen JJ, Cohn DL, Jasmer RM *et al.* An official ATS statement: hepatotoxicity of antituberculosis therapy. *Am. J. Respir. Crit. Care Med.*, 174(8), 935-952 (2006).
6. Wilby KJ, Ensom MH, Marra F. Review of evidence for measuring drug concentrations of first-line antitubercular agents in adults. *Clin. Pharmacokinet.*, 53(10), 873-890 (2014).
7. Perrin FM, Lipman MC, McHugh TD, Gillespie SH. Biomarkers of treatment response in clinical trials of novel antituberculosis agents. *Lancet Infect. Dis.*, 7(7), 481-490 (2007).
8. Nahid P, Saukkonen J, Mac Kenzie WR *et al.* CDC/NIH Workshop. Tuberculosis biomarker and surrogate endpoint research roadmap. *Am. J. Respir. Crit. Care Med.*, 184(8), 972-979 (2011).
   •• Summarizes recommendations for standardization of sample and data collection and creation of a specimen repository to support biomarker testing.
9. Gumbo T, Lenaerts AJ, Hanna D, Romero K, Nuermberger E. Nonclinical models for antituberculosis drug development: a landscape analysis. *J. Infect. Dis.*, 211 Suppl 3, S83-95 (2015).
   • Summarizes *in vitro* and animal models for the pathology and pharmacokinetic-pharmacodynamic aspects of human tuberculosis and outlines strengths and limitations of each.
10. Diacon AH, Donald PR. The early bactericidal activity of antituberculosis drugs. *Expert Rev. Anti Infect. Ther.*, 12(2), 223-237 (2014).
11. Almeida D, Nuermberger E, Tasneen R *et al.* Paradoxical effect of isoniazid on the activity of rifampin-pyrazinamide combination in a mouse model of tuberculosis. *Antimicrob. Agents Chemother.*, 53(10), 4178-4184 (2009).

12. Sirgel F, Venter A, Mitchison D. Sources of variation in studies of the early bactericidal activity of antituberculosis drugs. *J. Antimicrob. Chemother.*, 47(2), 177-182 (2001).
13. Mukamolova GV, Turapov O, Malkin J, Woltmann G, Barer MR. Resuscitation-promoting Factors Reveal an Occult Population of Tubercle Bacilli in Sputum. *Am. J. Respir. Crit. Care Med.*, 181(2), 174-180 (2010).
14. Bowness R, Boeree MJ, Aarnoutse R *et al.* The relationship between *Mycobacterium tuberculosis* MGIT time to positivity and cfu in sputum samples demonstrates changing bacterial phenotypes potentially reflecting the impact of chemotherapy on critical sub-populations. *J. Antimicrob. Chemother.*, 70(2), 448-455 (2015).
15. Wallis RS, Doherty TM, Onyebujoh P *et al.* Biomarkers for tuberculosis disease activity, cure, and relapse. *The Lancet infectious diseases*, 9(3), 162-172 (2009).
16. Phillips PP, Davies GR, Mitchison DA. Biomarkers for tuberculosis disease activity, cure, and relapse. *Lancet Infect. Dis.*, 10(2), 69-70; author reply 70-61 (2010).
17. Johnson JL, Hadad DJ, Dietze R *et al.* Shortening treatment in adults with noncavitary tuberculosis and 2-month culture conversion. *Am. J. Respir. Crit. Care Med.*, 180(6), 558-563 (2009).
18. Phillips PP, Nunn AJ, Paton NI. Is a 4-month regimen adequate to cure patients with non-cavitary tuberculosis and negative cultures at 2 months? *Int J Tuberc Lung Dis*, 17(6), 807-809 (2013).
19. Conde MB, Efron A, Loredo C *et al.* Moxifloxacin versus ethambutol in the initial treatment of tuberculosis: a double-blind, randomised, controlled phase II trial. *Lancet*, 373(9670), 1183-1189 (2009).
20. Burman WJ, Goldberg S, Johnson JL *et al.* Moxifloxacin versus ethambutol in the first 2 months of treatment for pulmonary tuberculosis. *Am. J. Respir. Crit. Care Med.*, 174(3), 331-338 (2006).
21. Gillespie SH, Crook AM, McHugh TD *et al.* Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis. *N. Engl. J. Med.*, 371(17), 1577-1587 (2014).
22. Jindani A, Harrison TS, Nunn AJ *et al.* High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis. *N. Engl. J. Med.*, 371(17), 1599-1608 (2014).
23. Jawahar MS, Banurekha VV, Paramasivan CN *et al.* Randomized Clinical Trial of Thrice-Weekly 4-Month Moxifloxacin or Gatifloxacin Containing Regimens in the Treatment of New Sputum Positive Pulmonary Tuberculosis Patients. *PLoS One*, 8(7) (2013).
24. Merle CS, Fielding K, Sow OB *et al.* A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis. *N. Engl. J. Med.*, 371(17), 1588-1598 (2014).
25. Wallis RS, Wang C, Meyer D, Thomas N. Month 2 culture status and treatment duration as predictors of tuberculosis relapse risk in a meta-regression model. *PLoS One*, 8(8), e71116 (2013).

26. Wallis RS, Peppard T, Hermann D. Month 2 culture status and treatment duration as predictors of recurrence in pulmonary tuberculosis: model validation and update. *PLoS One*, 10(4), e0125403 (2015).
•Meta-regression model of recurrence incorporating data from clinical studies, including recent fluroquinolone studies can inform the design of future phase 3 tuberculosis clinical trials based on 2-month culture status and regimen duration
27. Davies GR, Brindle R, Khoo SH, Aarons LJ. Use of nonlinear mixed-effects analysis for improved precision of early pharmacodynamic measures in tuberculosis treatment. *Antimicrob. Agents Chemother.*, 50(9), 3154-3156 (2006).
28. Sloan DJ, Mwandumba HC, Garton NJ *et al.* Pharmacodynamic Modeling of Bacillary Elimination Rates and Detection of Bacterial Lipid Bodies in Sputum to Predict and Understand Outcomes in Treatment of Pulmonary Tuberculosis. *Clin. Infect. Dis.*, 61(1), 1-8 (2015).
• Time to culture conversion in solid and liquid media was used to model bacillary elimination rate during first 8 weeks of treatment. Preliminary results demonstrated increased odds of unfavourable response with increase in % lipid body positive acid fast bacilli at 3-4 weeks of therapy.
29. Kurbatova EV, Cegielski JP, Lienhardt C *et al.* Sputum culture conversion as a prognostic marker for end-of-treatment outcome in patients with multidrug-resistant tuberculosis: a secondary analysis of data from two observational cohort studies. *Lancet Respir Med*, 3(3), 201-209 (2015).
30. Friedrich SO, Rachow A, Saathoff E *et al.* Assessment of the sensitivity and specificity of Xpert MTB/RIF assay as an early sputum biomarker of response to tuberculosis treatment. *Lancet Resp Med*, 1(6), 462-470 (2013).
31. Kim YJ, Lee SM, Park BK *et al.* Evaluation of propidium monoazide real-time PCR for early detection of viable *Mycobacterium tuberculosis* in clinical respiratory specimens. *Ann. Lab. Med.*, 34(3), 203-209 (2014).
32. Desjardin LE, Perkins MD, Wolski K *et al.* Measurement of sputum *Mycobacterium tuberculosis* messenger RNA as a surrogate for response to chemotherapy. *Am. J. Respir. Crit. Care Med.*, 160(1), 203-210 (1999).
33. Li L, Mahan CS, Palaci M *et al.* Sputum *Mycobacterium tuberculosis* mRNA as a marker of bacteriologic clearance in response to antituberculosis therapy. *J. Clin. Microbiol.*, 48(1), 46-51 (2010).
34. Honeyborne I, Mtafya B, Phillips PP *et al.* The molecular bacterial load assay replaces solid culture for measuring early bactericidal response to antituberculosis treatment. *J. Clin. Microbiol.*, 52(8), 3064-3067 (2014).
35. Honeyborne I, McHugh TD, Phillips PP *et al.* Molecular bacterial load assay, a culture-free biomarker for rapid and accurate quantification of sputum *Mycobacterium tuberculosis* bacillary load during treatment. *J. Clin. Microbiol.*, 49(11), 3905-3911 (2011).
36. Hamid Salim A, Aung KJ, Hossain MA, Van Deun A. Early and rapid microscopy-based diagnosis of true treatment failure and MDR-TB. *Int J Tuberc Lung Dis*, 10(11), 1248-1254 (2006).

37. Datta S, Sherman JM, Bravard MA, Valencia T, Gilman RH, Evans CA. Clinical evaluation of tuberculosis viability microscopy for assessing treatment response. *Clin. Infect. Dis.*, 60(8), 1186-1195 (2015).
38. Daniel J, Maamar H, Deb C, Sirakova TD, Kolattukudy PE. *Mycobacterium tuberculosis* uses host triacylglycerol to accumulate lipid droplets and acquires a dormancy-like phenotype in lipid-loaded macrophages. *PLoS Pathog.*, 7(6), e1002093 (2011).
39. Kayigire XA, Friedrich SO, van der Merwe L, Donald PR, Diacon AH. Simultaneous staining of sputum smears for acid-fast and lipid-containing *Myobacterium tuberculosis* can enhance the clinical evaluation of antituberculosis treatments. *Tuberculosis (Edinb)*, (2015).
40. Wallis RS, Vinhas SA, Johnson JL *et al.* Whole blood bactericidal activity during treatment of pulmonary tuberculosis. *J. Infect. Dis.*, 187(2), 270-278 (2003).
41. Wallis RS, Jakubiec W, Mitton-Fry M *et al.* Rapid Evaluation in Whole Blood Culture of Regimens for XDR-TB Containing PNU-100480 (Sutezolid), TMC207, PA-824, SQ109, and Pyrazinamide. *PLoS One*, 7(1) (2012).
42. Wallis RS, Perkins M, Phillips M *et al.* Induction of the antigen 85 complex of *Mycobacterium tuberculosis* in sputum: a determinant of outcome in pulmonary tuberculosis treatment. *J. Infect. Dis.*, 178(4), 1115-1121 (1998).
43. Wallis RS, Phillips M, Johnson JL *et al.* Inhibition of isoniazid-induced expression of *Mycobacterium tuberculosis* antigen 85 in sputum: potential surrogate marker in tuberculosis chemotherapy trials. *Antimicrob. Agents Chemother.*, 45(4), 1302-1304 (2001).
44. Wood R, Racow K, Bekker LG *et al.* Lipoarabinomannan in urine during tuberculosis treatment: association with host and pathogen factors and mycobacteriuria. *BMC Infect. Dis.*, 12, 47 (2012).
45. Drain PK, Gounder L, Grobler A, Sahid F, Bassett IV, Moosa MY. Urine lipoarabinomannan to monitor antituberculosis therapy response and predict mortality in an HIV-endemic region: a prospective cohort study. *BMJ Open*, 5(4), e006833 (2015).
46. Young BL, Mlamla Z, Gqamana PP *et al.* The identification of tuberculosis biomarkers in human urine samples. *Eur. Respir. J.*, 43(6), 1719-1729 (2014).
47. Ehrlich RI, Adams S, Baatjies R, Jeebhay MF. Chronic airflow obstruction and respiratory symptoms following tuberculosis: a review of South African studies. *Int J Tuberc Lung Dis*, 15(7), 886-891 (2011).
48. Plit ML, Anderson R, Van Rensburg CE *et al.* Influence of antimicrobial chemotherapy on spirometric parameters and pro-inflammatory indices in severe pulmonary tuberculosis. *Eur. Respir. J.*, 12(2), 351-356 (1998).
49. Hnizdo E, Singh T, Churchyard G. Chronic pulmonary function impairment caused by initial and recurrent pulmonary tuberculosis following treatment. *Thorax*, 55(1), 32-38 (2000).
50. Ralph AP, Ardian M, Wiguna A *et al.* A simple, valid, numerical score for grading chest x-ray severity in adult smear-positive pulmonary tuberculosis. *Thorax*, 65(10), 863-869 (2010).

51. Perrin FM, Woodward N, Phillips PP *et al.* Radiological cavitation, sputum mycobacterial load and treatment response in pulmonary tuberculosis. *Int J Tuberc Lung Dis*, 14(12), 1596-1602 (2010).
52. Pefura-Yone EW, Kuaban C, Assamba-Mpom SA, Moifo B, Kengne AP. Derivation, validation and comparative performance of a simplified chest X-ray score for assessing the severity and outcome of pulmonary tuberculosis. *Clin. Respir. J.*, 9(2), 157-164 (2015).
53. Kriel M, Lotz JW, Kidd M, Walzl G. Evaluation of a radiological severity score to predict treatment outcome in adults with pulmonary tuberculosis. *Int J Tuberc Lung Dis*, 19(11), 1354-1360 (2015).
54. Hamilton CD, Stout JE, Goodman PC *et al.* The value of end-of-treatment chest radiograph in predicting pulmonary tuberculosis relapse. *Int J Tuberc Lung Dis*, 12(9), 1059-1064 (2008).
    - Simple validated tool for grading chest x-ray severity which predicts sputum conversion and may have utility across different settings.
55. Bosch-Marcet J, Serres-Creixams X, Zuasnabar-Cotro A, Codina-Puig X, Catala-Puigbo M, Simon-Riazuelo JL. Comparison of ultrasound with plain radiography and CT for the detection of mediastinal lymphadenopathy in children with tuberculosis. *Pediatr. Radiol.*, 34(11), 895-900 (2004).
56. Bosch-Marcet J, Serres-Creixams X, Borras-Perez V, Coll-Sibina MT, Guitet-Julia M, Coll-Rosell E. Value of sonography for follow-up of mediastinal lymphadenopathy in children with tuberculosis. *J. Clin. Ultrasound*, 35(3), 118-124 (2007).
57. Sharma SK, Solanki R, Mohan A, Jain NK, Chauhan LS, Pleural Effusion Study G. Outcomes of Category III DOTS treatment in immunocompetent patients with tuberculosis pleural effusion. *Int J Tuberc Lung Dis*, 16(11), 1505-1509 (2012).
58. Jain R, Sawhney S, Bhargava DK, Berry M. Diagnosis of abdominal tuberculosis: sonographic findings in patients with early disease. *AJR Am. J. Roentgenol.*, 165(6), 1391-1395 (1995).
59. Huang WC, Tseng CW, Chang KM, Hsu JY, Chen JH, Shen GH. Usefulness of tumor marker CA-125 serum levels for the follow-up of therapeutic responses in tuberculosis patients with and without serositis. *Jpn. J. Infect. Dis.*, 64(5), 367-372 (2011).
60. Chen RY, Dodd LE, Lee M *et al.* PET/CT imaging correlates with treatment outcome in patients with multidrug-resistant tuberculosis. *Sci. Transl. Med.*, 6(265), 265ra166 (2014).
61. Coleman MT, Chen RY, Lee M *et al.* PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Sci. Transl. Med.*, 6(265), 265ra167 (2014).
62. Ghesani N, Patrawalla A, Lardizabal A, Salgame P, Fennelly KP. Increased cellular activity in thoracic lymph nodes in early human latent tuberculosis infection. *Am. J. Respir. Crit. Care Med.*, 189(6), 748-750 (2014).
63. Lawn SD, Obeng J, Acheampong JW, Griffin GE. Resolution of the acute-phase response in West African patients receiving treatment for pulmonary tuberculosis. *Int J Tuberc Lung Dis*, 4(4), 340-344 (2000).

64. Bajaj G, Rattan A, Ahmad P. Prognostic value of 'C' reactive protein in tuberculosis. *Indian Pediatr.*, 26(10), 1010-1013 (1989).
65. Scott GM, Murphy PG, Gemidjioglu ME. Predicting deterioration of treated tuberculosis by corticosteroid reserve and C-reactive protein. *J. Infect.*, 21(1), 61-69 (1990).
    • This study demonstrated, both in macaques and humans, that PET-CT may have use as a quantitative measure of drug efficacy and treatment response as early as 1 month.
66. Ugajin M, Miwa S, Shirai M *et al.* Usefulness of serum procalcitonin levels in pulmonary tuberculosis. *Eur. Respir. J.*, 37(2), 371-375 (2011).
67. Clifford V, Zufferey C, Street A, Denholm J, Tebruegge M, Curtis N. Cytokines for monitoring anti-tuberculous therapy: A systematic review. *Tuberculosis (Edinb)*, 95(3), 217-228 (2015).
68. Djoba Siawaya JF, Beyers N, van Helden P, Walzl G. Differential cytokine secretion and early treatment response in patients with pulmonary tuberculosis. *Clin. Exp. Immunol.*, 156(1), 69-77 (2009).
69. Mihret A, Abebe M, Bekele Y, Aseffa A, Walzl G, Howe R. Impact of HIV co-infection on plasma level of cytokines and chemokines of pulmonary tuberculosis patients. *BMC Infect. Dis.*, 14, 125 (2014).
70. Harari A, Rozot V, Bellutti Enders F *et al.* Dominant TNF-alpha+ *Mycobacterium tuberculosis*-specific CD4+ T cell responses discriminate between latent infection and active disease. *Nat. Med.*, 17(3), 372-376 (2011).
71. Day CL, Abrahams DA, Lerumo L *et al.* Functional capacity of *Mycobacterium tuberculosis*-specific T cell responses in humans is associated with mycobacterial load. *J. Immunol.*, 187(5), 2222-2232 (2011).
72. Millington KA, Innes JA, Hackforth S *et al.* Dynamic relationship between IFN-gamma and IL-2 profile of *Mycobacterium tuberculosis*-specific T cells and antigen load. *J. Immunol.*, 178(8), 5217-5226 (2007).
73. Feruglio SL, Tonby K, Kvale D, Dyrhol-Riise AM. Early dynamics of T helper cell cytokines and T regulatory cells in response to treatment of active *Mycobacterium tuberculosis* infection. *Clin. Exp. Immunol.*, 179(3), 454-465 (2015).
74. Essone PN, Kalsdorf B, Chegou NN *et al.* Bifunctional T-cell-derived cytokines for the diagnosis of tuberculosis and treatment monitoring. *Respiration*, 88(3), 251-261 (2014).
75. Parkash O, Agrawal S, Madhan Kumar M. T regulatory cells: Achilles' heel of *Mycobacterium tuberculosis* infection? *Immunol. Res.*, 62(3), 386-398 (2015).
    • This review addresses the role of T regulatory lymphocytes (T_{regs}) in disease pathogenesis in TB and discusses monitoring of T_{regs}) as a potential biomarker of treatment response in both pulmonary and extrapulmonary TB.
76. Adekambi T, Ibegbu CC, Cagle S *et al.* Biomarkers on patient T cells diagnose active tuberculosis and monitor treatment response. *J. Clin. Invest.*, 125(5), 1827-1838 (2015).

77. Rabna P, Andersen A, Wejse C *et al.* Utility of the plasma level of suPAR in monitoring risk of mortality during TB treatment. *PLoS One*, 7(8), e43933 (2012).
78. Ong CW, Elkington PT, Friedland JS. Tuberculosis, pulmonary cavitation, and matrix metalloproteinases. *Am. J. Respir. Crit. Care Med.*, 190(1), 9-18 (2014).
79. Elkington P, Shiomi T, Breen R *et al.* MMP-1 drives immunopathology in human tuberculosis and transgenic mice. *J Clin Invest*, 121(5), 1827-1833 (2011).
80. Walker NF, Clark SO, Oni T *et al.* Doxycycline and HIV infection suppress tuberculosis-induced matrix metalloproteinases. *Am. J. Respir. Crit. Care Med.*, 185(9), 989-997 (2012).
81. Ugarte-Gil CA, Elkington P, Gilman RH *et al.* Induced sputum MMP-1, -3 & -8 concentrations during treatment of tuberculosis. *PLoS One*, 8(4), e61333 (2013).
82. Coussens AK, Wilkinson RJ, Nikolayevskyy V *et al.* Ethnic variation in inflammatory profile in tuberculosis. *PLoS Pathog.*, 9(7), e1003468 (2013).
83. Thwaites GE, Simmons CP, Than Ha Quyen N *et al.* Pathophysiology and prognosis in vietnamese adults with tuberculous meningitis. *J. Infect. Dis.*, 188(8), 1105-1115 (2003).
84. Andrade BB, Pavan Kumar N, Amaral EP *et al.* Heme Oxygenase-1 Regulation of Matrix Metalloproteinase-1 Expression Underlies Distinct Disease Profiles in Tuberculosis. *J. Immunol.*, 195(6), 2763-2773 (2015).
85. Berry MP, Graham CM, McNab FW *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*, 466(7309), 973-977 (2010).
•• Highly cited study defining the blood transcriptomic signature of active tuberculosis which correlate with disease severity and reverting in response to treatment.
86. Bloom CI, Graham CM, Berry MP *et al.* Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PLoS One*, 7(10), e46191 (2012).
87. Cliff JM, Lee JS, Constantinou N *et al.* Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. *J. Infect. Dis.*, 207(1), 18-29 (2013).
88. Cliff JM, Cho JE, Lee JS *et al.* Excessive Cytolytic Responses Predict Tuberculosis Relapse After Apparently Successful Treatment. *J. Infect. Dis.*, 213(3), 485-495 (2016).
89. Wilkinson RJ, Vordermeier HM, Wilkinson KA *et al.* Peptide-specific T cell response to *Mycobacterium tuberculosis*: clinical spectrum, compartmentalization, and effect of chemotherapy. *J. Infect. Dis.*, 178(3), 760-768 (1998).
90. Clifford V, He Y, Zufferey C, Connell T, Curtis N. Interferon gamma release assays for monitoring the response to treatment for tuberculosis: A systematic review. *Tuberculosis (Edinb)*, 95(6), 639-650 (2015).

91. Wilkinson KA, Kon OM, Newton SM *et al.* Effect of treatment of latent tuberculosis infection on the T cell response to *Mycobacterium tuberculosis* antigens. *J. Infect. Dis.*, 193(3), 354-359 (2006).
92. Dyrhol-Riise AM, Gran G, Wentzel-Larsen T, Blomberg B, Haanshuus CG, Morkve O. Diagnosis and follow-up of treatment of latent tuberculosis; the utility of the QuantiFERON-TB Gold In-tube assay in outpatients from a tuberculosis low-endemic country. *BMC Infect. Dis.*, 10, 57 (2010).
93. Torres M, Garcia-Garcia L, Cruz-Hervert P *et al.* Effect of isoniazid on antigen-specific interferon-gamma secretion in latent tuberculosis. *Eur. Respir. J.*, 45(2), 473-482 (2015).
94. Nicol MP, Gnanashanmugam D, Browning R *et al.* A Blueprint to Address Research Gaps in the Development of Biomarkers for Pediatric Tuberculosis. *Clin. Infect. Dis.*, 61Suppl 3, S164-172 (2015).
95. Wallis RS, Peppard T. Early Biomarkers and Regulatory Innovation in Multidrug-Resistant Tuberculosis. *Clin. Infect. Dis.*, 61Suppl 3, S160-163 (2015).
96. Fol M, Druszczynska M, Wlodarczyk M, Ograczyk E, Rudnicka W. Immune response gene polymorphisms in tuberculosis. *Acta Biochim. Pol.*, (2015).
97. Nahid P, Bliven-Sizemore E, Jarlsberg LG *et al.* Aptamer-based proteomic signature of intensive phase treatment response in pulmonary tuberculosis. *Tuberculosis (Edinb)*, 94(3), 187-196 (2014).
98. Moller M, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. *FEMS Immunol. Med. Microbiol.*, 58(1), 3-26 (2010).
99. Esterhuyse MM, Linhart HG, Kaufmann SH. Can the battle against tuberculosis gain from epigenetic research? *Trends Microbiol.*, 20(5), 220-226 (2012).
100. Prideaux B, Via LE, Zimmerman MD *et al.* The association between sterilizing activity and drug distribution into tuberculosis lesions. *Nat. Med.*, 21(10), 1223-1227 (2015).
•• ŸMass spectrometry imaging was used to demonstrate differences in spatio-temporal distribution of rifampicin, pyrazinamide and moxifloxacin in caseous foci and cellular levels in resected lung sections.
101. Pienaar E, Dartois V, Linderman JJ, Kirschner DE. In silico evaluation and exploration of antibiotic tuberculosis treatment regimens. *BMC Syst. Biol.*, 9, 79 (2015).
- A computational tool comprising numerous factors including immunological factors, drug penetration at the level of the granuloma and host pharmacokinetics and –dynamics to assist in the evaluation of treatment response.

Type of file: table
Label: Tables 1-4
Filename: Tabs1-4.docx

Table 1 Treatment response and outcomes in active TB (adapted from [4])

Treatment response/outcome	Drug susceptible TB	Drug resistant TB
Interim		
Paradoxical reaction	Paradoxical worsening or recurring of tuberculous lesions or development of new lesions despite successful anti-TB treatment in response to an exaggerated and dyregulated immune response in the context of rapidly recovering immunity and presence of abundant infective antigen	
Culture conversion	Two month culture conversion: Having shown baseline culture positivity, at least 1 negative culture by 2 months	Having shown baseline culture positivity, 2 consecutive cultures, taken at least 30 days apart are found to be negative
Culture reversion		After initial conversion, 2 consecutive cultures, taken at least 30 days apart, are found to be positive.
Long term		
Treatment cure	A patient with bacteriologically confirmed TB at the beginning of treatment who was smear- or culture-negative in the last month of treatment and on at least one previous occasion	A patient with bacteriologically confirmed TB who has completed treatment as recommended by the national policy without evidence of failure and 3 or more consecutive negative cultures taken at least 30 days apart post completion of the intensive phase
Treatment completer	A TB patient who completed treatment without evidence of failure but with no record to show that sputum smear or culture results in the last month of treatment and on at least one previous occasion were negative	Treatment completed as recommended by the national policy without evidence of failure but no record of 3 or more negative consecutive cultures taken at least 30 days apart post completion of the intensive phase
Treatment success	The sum of cured and treatment completed	
Treatment failure	A TB patient whose sputum smear or culture is positive at month 5 or later during treatment.	Treatment terminated or need for permanent regimen change of at least 2 anti-TB drugs because any of: − lack of conversion by the end of the intensive phase − bacteriological reversion in the continuation phase after conversion to negative − evidence of acquired drug resistance − adverse drug reactions/paradoxical reaction
Treatment relapse	Patients who were declared cured or treatment completed at the end of their most recent course of TB treatment, and are now diagnosed with a recurrent episode of TB. This can be either a true relapse or a new episode of TB caused by reinfection*	
Lost to follow up	A TB patient who did not start treatment or whose treatment was interrupted for 2 consecutive months or more	
Died	A TB patient who dies for any reason before starting or during the course of treatment	

Table 2 Sputum (direct specimen or cultured isolate) based methodology for treatment monitoring

Methodology	Utiliity	Limitations	Ref
Early bactericidal activity as measured by rate of fall in CFUs/increased time to positivity in liquid culture over first 2-14 days of treatment	Bactericidal efficacy (including synergism/antagonism) of single drug or regimen and its/their relationship to PK can be ascertained. Results can be ascertained relatively quickly and therefore guide progression to phase 2 clinical trials	There can be variability between different centers and intra-individual variation pre-treatment. Results can be affected by media and growth factors used. Sterilizing activity against non-replicating persisters is not accurately predicted	[12,14]
2 month sputum culture conversion and time to culture conversion.	Significant correlation with combined rate of failure, relapse and death and provides an intermediate bacteriological endpoint. Mixed effects modeling of time to culture conversion data in solid and liquid media can better characterize pharmacodynamics of treatment	In both DS and DR TB, neither 2-month culture conversion as a binary outcome nor time to culture conversion is a perfect surrogate end-point when considering long term outcomes in the individual and as a trial efficacy endpoint. Inter-laboratory variation in SOPs can also affect comparability results.	[15-18,25,26,28,29]
Polymerase chain reaction based methods	Provides real time quantification of bacterial load pre-treatment and to a degree, during drug treatment.	DNA/RNA amplification from dead bacteria can confound ascertainment of sputum sterilization during treatment. The clinical significance of detection of DNA/RNA, in the context of non-culturability at different time points during treatment is unclear.	[30,32-35]
Viability and lipid body staining of sputum	Assessment of viability via staining with fluorescein diacetate may expedite diagnosis of poor response to treatment and quantify early response to treatment. Change in percentage rise in % lipid body positive AFB over the first 4 weeks may predict failure/relapse.	These assays are not applicable to AFB negative sputa. Serial samples taken longitudinally during treatment are required for assessment and a single sample e.g. pre-treatment has limited utility.	[28,36,37]
Whole blood bactericidal activity	Inferior bactericidal activity seen in whole blood cultures of sputa taken from patients at an interim time point in treatment may predict delayed sputum sterilization/subsequent relapse	The assay may be influenced by factors unrelated to outcome e.g. sputum volume and viscosity. This is a culture based technique and can take several weeks.	[40,41]

Abbreviations: AFB acid fast bacilli, CFU colony forming unit, DNA deoxyribonucleic acid, RNA ribonucleic acid, DS drug sensitive, DR drug resistant, TB tuberculosis, PK pharmacokinetics, SOP standard operating procedure.

Table 3 Radiological based methodology for treatment monitoring

Methodology	Utility	Limitations	Ref
Chest radiograph	The Timika score has been validated in 3 geographical populations (1 cohort 30% HIV co-infected) and has moderate sensitivity and specificity for predicting smear non-conversion at 2 months. Persistent cavity after 6 months of TB treatment was independently associated with disease relapse	Inter-observer variability Not proven to accurately predict failure/relapse	[50-54]
Ultrasound	Paediatric: High sensitivity for diagnosis of active intrathoracic TB and can detect response to treatment inside 3 months.	Paediatric: Cannot distinguish active TB from other causes of mediastinal lymphadenopathy. No evidence for utility in HIV co-infection.	[55,56]
	Pleural: Can provide confirmation of cure at 6 months in around 90% of cases of small, unilateral effusion. May also confirm response by 2 months.	Misses approximately 10% of those who are truly cured at 6 months. No evidence for utility in HIV co-infection	[57,59]
	Peritoneal: Can provide confirmation of cure at 6 months in cases with initial mesenteric thickness >15mm. May also confirm response by 2 months.	Peritoneal: Less reliable in cases without significant mesenteric thickening. No evidence for utility in HIV co-infection.	[58,59]
PET/CT	Active TB: Appears promising as early outcome measure (greater accuracy at predicting durable cure than sputum culture or CT alone) Appears promising for comparative drug trials Applicable to macaque studies Latent TB: Has potential for monitoring treatment response May be applicable to HIV co-infection	Active TB: Highly expensive purchase and running costs High level radiation exposure Larger numbers required for statistically significant benefit Latent TB: Larger numbers required to distinguish treatment response from natural history.	[60-62]

Table 4 Immune markers for treatment monitoring

Immune marker	Utility	Limitations	Ref
<u>Acute phase proteins:</u> C-reactive protein (CRP)	-Fall in CRP (to within normal range) in response to treatment correlates with therapeutic response. A threshold of 20mg/l at baseline, which persisted during treatment has been associated with adverse outcome.	-Not specific to tuberculosis	[63-65]
Procalcitonin (PCT)	- PCT levels ≥0.5 ng/ml^{-1} maybe indicative of poor prognosis and increased mortality risk	- Not specific to tuberculosis and mostly applicable only to disseminated tuberculosis	[66]
<u>Cytokines and immune activation markers :</u> Tumour necrosis factor (TNF) α, interleukin-2 (IL-2) and interferon (IFN) Y	-Decreased TNFα during treatment has been seen in both TB-antigen stimulated and un-stimulated cytokine response. During successful treatment *M. tuberculosis* (MTB) specific T cells change from single positive TNFα expressing phenotype (related to bacterial load) to either poly-functional MTB-specific T cell profile (expressing IFN-γ^+ , TNFα^+ and IL-2^+) or IL-2^+ single producing and/or IL-2^+ and IFN-γ^+ double producing MTB-specific T cell profile.	-No specific thresholds specified	[70-74,76]
Soluble urokinase plasminogen activator receptor (suPAR)	-High plasma levels of suPAR at inclusion or at any point during treatment was associated with increased mortality.	-No specific thresholds specified and not specific to tuberculosis	[77]
matrix metalloproteinases (MMPs)	MMP-1, -3 and -8 concentrations in sputum decline in response to successful ATT and delayed culture conversion is positively correlated with MMP-1.	Variation in production, regulation and biological activity of MMPs in relation to different points in disease pathogenesis. Significant ethnic heterogeneity in MMP profile during anti-tuberculosis therapy	[78,80-84]