Primary efficacy endpoints in phase 3 non-inferiority trials to establish new tuberculosis treatment regimens should only include microbiological outcomes

Lancet Microbe 2025
Published Online
https://doi.org/10.1016/j.lanmic.2025.101117

Treatment regimens for tuberculosis remain protracted and burdensome, both for patients to adhere to and for health-care providers to administer. Particularly, many current regimens for drug-resistant tuberculosis are poorly tolerated by patients and carry an unacceptable side-effect profile. The rapid emergence of bedaquiline resistance necessitates treatment regimens that are active against resistant strains, regimens with a high genetic barrier to acquired resistance, and regimens containing antimicrobials for which near-patient drug-susceptibility testing is available. A pressing need therefore exists to develop and evaluate novel tuberculosis treatment regimens.

The relative efficacy and safety of tuberculosis treatment regimens can only be reliably ascertained through randomised controlled trials. However, trial capacity for tuberculosis is severely constrained. Tuberculosis trial capacity should be urgently expanded. Furthermore, it is crucial that existing capacity is used well, warranting improvements in statistical and operational efficiency[1] and trials designed to produce unbiased estimates of efficacy and safety that can inform policy and clinical decision making in communities experiencing a high burden of tuberculosis disease. The choice of primary efficacy endpoints impacts both statistical efficiency and ability to produce unbiased results that have relevance beyond patients enrolled in clinical trials.

Tuberculosis trials typically adopt a non-inferiority design for the primary efficacy endpoint, while aiming to show superiority in other domains, such as quality of life or treatment duration. This approach is used because the current standard of care is highly efficacious under trial conditions. Treatment regimens for drug-resistant tuberculosis have also achieved very low rates of microbiological failure or relapse in trials.[2,3]

Composite primary efficacy endpoints have typically included treatment failure, relapse, death, early discontinuation of treatment, loss to follow-up, and treatment switches, with each of these components carrying equal weight. In this Comment, we present four arguments in favour of abandoning this approach for a primary efficacy endpoint comprising only microbiological events—treatment failure and relapse.

First, giving equal weight in a composite primary outcome to microbiological endpoints and soft outcomes, such as stopping treatment a few weeks early or having to change components of a tuberculosis treatment regimen, makes little sense. Many people who stop treatment early experience no ill consequences. If stopping treatment does lead to treatment failure or relapse, these outcomes will anyway be captured as microbiological endpoints. Information on losses to follow-up, adverse events, tolerability, treatment adherence, and treatment changes can still be captured. Principled approaches to handling treatment switches, including prespecified strategies and re-randomisations, are outlined elsewhere.[1,4] A pragmatic approach to defining microbiological endpoints for patients who die, are lost to follow-up, or cannot produce sputum has been proposed by Phillips and colleagues.[3,5]

Second, given that rates of loss to follow-up are much higher in routine practice than in randomised controlled trials,[2,6–11] including loss to follow-up in a composite primary efficacy endpoint will produce estimates of efficacy that cannot be generalised to settings in which most people with tuberculosis receive treatment. Some loss to follow-up can be attributed to treatment regimens, as poor tolerability can result in patients abandoning treatment. Poor tolerability can be particularly problematic when treatment programmes do not have the resources needed to offer patients antiemetics and other forms of symptomatic relief. Programmatic factors, such as clinics being located far from patients' homes, can also contribute to loss to follow-up. Phase 3 trials seeking to establish new treatment regimens should capture data on tolerability, treatment adherence, and loss to follow-up. Loss to follow-up is influenced by many factors beyond the choice of treatment regimen, and these factors might not generalise to non-trial settings; therefore, loss to follow-up should not be included as part of the primary efficacy outcome.

Third, when regimens being tested are of different durations, using an endpoint that includes early cessation of treatment or treatment switches will bias estimates of efficacy in favour of the shorter regimen, because patients on longer regimens will have more opportunity to stop or change treatment compared with those on shorter

	Number of participants (mITT)	Treatment failure or relapse (%)	Other events (%)	NI margin (%)	Risk difference per cent (95% CI)		Percentage reduction in width of CI (%)
					All outcomes	Microbiological outcomes only	
STREAM 1[6]							
Standard of care	124	7 (6)	18 (15)	10	Reference category		
9–11 months including injectables	245	26 (11)	26 (11)		1·1 (–7·7 to 9·8)	5·0 (–0·6 to 10·6)	36
NExT[7]							
Standard of care	44	6 (14)	28 (64)	Not provided	Reference category		
6–9 months all oral	49	4 (8)	20 (41)		–28·3 (–47·0 to –9·6)	–5·5 (–18·2 to 7·2)	32
ZeNix[8]							
Lzd 1200 mg 26 weeks	44	0	3 (7)	Not provided	Reference category*		
Lzd 1200 mg 9 weeks	45	2 (4)	3 (7)		4·3 (–7·5 to 16·1)	4·4 (–1·6 to 10·5)	49
Lzd 600 mg 26 weeks	45	1 (2)	3 (7)		2·1 (–9·1 to 13·2)	2·2 (–2·1 to 6·5)	61
Lzd 600 mg 9 weeks	44	2 (5)	5 (11)		9·1 (–4·0 to 22·2)	4·5 (–1·6 to 10·7)	53
TB PRACTECAL[9]							
Standard of care	137	0	56 (41)	12	Reference category		
BPaLzdM	138	1 (1)	15 (11)		–29·3 (–39·1 to –19·5)	0·7 (–0·7 to 2·1)	86
BPaLzdC	115	6 (5)	10 (9)		–27·0 (–37·3 to –16·6)	5·2 (1·2 to 9·3)	61
BPaLzd	111	3 (3)	12 (11)		–27·4 (–37·8 to –17·0)	2·7 (–0·3 to 5·7)	71
STREAM 2[2]							
Standard of care 1	187	20 (11)	34 (18)	10	Reference category		
9 months all oral	196	8 (4)	26 (12)		–11·5 (–19·9 to –3·1)	–6·6 (–11·8 to –1·4)	38
Standard of care 2	127	16 (13)	24 (19)		Reference category		
6 months including injectables	134	3 (2)	9 (7)		–22·5 (–32·0 to –13·1)	–10·4 (–16·7 to –4·1)	33
endTB[10]							
Standard of care	119	1 (1)	22 (18)	12	Reference category		
BLzdMZ	118	1 (1)	12 (10)		–8·3 (–17·4 to 0·8)	0·0 (–2·3 to 2·3)	74
BCLzdLfxZ	115	3 (3)	8 (7)		–9·8 (–18·7 to –0·9)	1·8 (–1·6 to 5·1)	62
BDLzdLfxZ	122	4 (3)	14 (11)		–4·6 (–14·1 to 4·9)	2·4 (–1·1 to 6·0)	62
DCLzdLfxZ	118	13 (11)	12 (10)		1·9 (–8·4 to 12·1)	10·2 (4·3 to 16·1)	43
DCMZ	107	10 (9)	8 (7)		–2·5 (–12·5 to 7·5)	8·5 (2·8 to 14·3)	43

B=bedaquiline. C=clofazimine. D=delamanid. Lfx=levofloxacin. Lzd=linezolid. mITT=modified intention-to-treat. M=moxifloxacin. NI=non-inferiority. Pa=pretomamid. Z=pyrazinamide. *This trial was not designed to provide pairwise comparisons between groups and, as such, no control group was specified.

Table: **Difference between interventional regimens and comparator in recent trials in drug-resistant tuberculosis, calculated using all components of the composite primary efficacy outcome and only using microbiological events**

regimens. Some examples of this phenomenon are given in a subsequent paragraph.

Finally, had recent tuberculosis treatment trials adopted a primary efficacy endpoint comprising only microbiological events (treatment failure or relapse), the trials would have had increased power or required fewer participants, because the standard error of a risk difference is a function of both the number of patients in each arm (n1 and n0) and the proportions of participants who experience the event in each arm (p1 and p0). The quantity p(1–p) is maximised at a probability of 0·5 and minimised at the limits; consequently, precision improves as outcome probabilities tend towards 0 or 1.

$$S.E\left(p1-p0\right)=\sqrt{\left[\frac{p1\left(1-p1\right)}{n1}+\frac{p0\left(1-p0\right)}{n0}\right]}$$

To show the improved precision that might be achieved, we used data from published randomised controlled trials of treatment regimens for drug-resistant tuberculosis.[2,6–10] The population size, number of events included in the composite primary outcome, non-inferiority margin, and two estimates of the primary efficacy outcome, with their associated confidence intervals, are presented (table). We present one estimate, calculated by us, using the primary outcome as defined by the authors and one estimate, also calculated by us, including only microbiological endpoints (treatment failure or relapse). The final column notes the improvement in precision achieved by switching to a purely microbiological primary endpoint, given as the percentage reduction in the width of the confidence interval. These estimates of relative efficacy might differ somewhat from those included in the published papers, as

some improvement in precision could have been achieved by adjusting for baseline covariates and randomisation stratification factors.

In most trials, non-microbiological endpoints were much more frequent than microbiological endpoints. This observation was particularly striking with longer standard of care regimens. For example, in the standard of care arm in the TB-PRACTECAL trial, no microbiological failures or relapses occurred among 56 events included in the composite primary efficacy endpoint. As expected, precision was markedly improved using only microbiological endpoints, with a 32–86% reduction in the width of the confidence intervals. Of note, despite the confidence intervals narrowing, two intervention regimens that met the prespecified non-inferiority criteria using a composite primary efficacy endpoint—the intervention regimen in the STREAM 1 trial and the DCMZ regimen in the endTB trial—failed to do so with a microbiological primary endpoint. In both instances, in the longer standard of care regimen, a large number of non-microbiological endpoints was observed, while in the intervention arms, treatment failure or relapse occurred more frequently.

Potential objections to our proposed approach include omitting death from the primary efficacy outcome and the challenges in determining microbiological endpoints in trials.

Death is certainly an outcome that matters to people with tuberculosis. However, in tuberculosis trials, most deaths are not attributable to tuberculosis. By including events in a composite primary outcome that cannot be impacted by the intervention, there is a risk of falsely concluding non-inferiority. While an endpoint review committee can be formed, which is masked to treatment allocation, to assess whether deaths are caused by tuberculosis, our preference would be to include death as a safety outcome.

Ascertaining microbiological outcomes in tuberculosis trials is often challenging, particularly when treatment stops a productive cough. Potential solutions include sputum induction or adopting a pragmatic definition of microbiological cure. A person who is well, with no signs or symptoms of tuberculosis disease, 6 months after completing treatment and who cannot produce sputum despite appropriate coaching on technique could be considered microbiologically negative. Importantly, including additional non-microbiological events in a composite primary efficacy endpoint does not help with the assessment of efficacy and, as we have explained, might make matters worse.

Restricting composite primary efficacy endpoints to include only microbiological events prioritises the most consequential treatment outcomes, produces an efficacy outcome that is relevant to routine clinical practice, reduces bias when treatment regimens are of different durations, and improves study power. This approach should be the default choice in definitive tuberculosis treatment trials.

We declare no competing interests.

*Tom A Yates, *Daniel J Grint*
daniel.grint@lshtm.ac.uk

Institute of Health Informatics, University College London, London, UK (TAY); Division of Infection and Immunity, University College London, London, UK (TAY); Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK (DJG)

1 Yates TA, Barnes S, Dedicoat M, et al. Drug-resistant tuberculosis treatments, the case for a phase III platform trial. *Bull World Health Organ* 2024; **102:** 657–64.
2 Goodall RL, Meredith SK, Nunn AJ, et al. Evaluation of two short standardised regimens for the treatment of rifampicin-resistant tuberculosis (STREAM stage 2): an open-label, multicentre, randomised, non-inferiority trial. *Lancet* 2022; **400:** 1858–68.
3 Phillips PPJ, Van Deun A, Ahmed S, et al. Investigation of the efficacy of the short regimen for rifampicin-resistant TB from the STREAM trial. *BMC Med* 2020; **18:** 314.
4 Pham TM, Tweed CD, Carpenter JR, et al. Rethinking intercurrent events in defining estimands for tuberculosis trials. *Clin Trials* 2022; **19:** 522–33.
5 Phillips PPJ, Weir IR, Dufault SM, Hills N. Modernization of endpoints and estimands of late-phase tuberculosis therapeutic trials. Jan 13, 2024. https://osf.io/4a7cq/ (accessed March 5, 2025).
6 Nunn AJ, Phillips PPJ, Meredith SK, et al. A trial of a shorter regimen for rifampin-resistant tuberculosis. *N Engl J Med* 2019; **380:** 1201–13.
7 Esmail A, Oelofse S, Lombard C, et al. An all-oral 6-month regimen for multidrug-resistant tuberculosis: a multicenter, randomized controlled clinical trial (the NExT Study). *Am J Respir Crit Care Med* 2022; **205:** 1214–27.
8 Conradie F, Bagdasaryan TR, Borisov S, et al. Bedaquiline–pretomanid–linezolid regimens for drug-resistant tuberculosis. *N Engl J Med* 2022; **387:** 810–23.
9 Nyang'wa BT, Berry C, Kazounis E, et al. Short oral regimens for pulmonary rifampicin-resistant tuberculosis (TB-PRACTECAL): an open-label, randomised, controlled, phase 2B–3, multi-arm, multicentre, non-inferiority trial. *Lancet Respir Med* 2024; **12:** 117–28.
10 Guglielmetti L, Khan U, Velásquez GE, et al. Oral regimens for rifampin-resistant, fluoroquinolone-susceptible tuberculosis. *N Engl J Med* 2025; **392:** 468–82.
11 WHO. Global tuberculosis report 2024. Oct 29, 2024. https://www.who.int/publications/i/item/9789240101531 (accessed March 5, 2025).