

A Standardized Approach for Collection of Objective Data to Support Outcome Determination for Late-Phase Tuberculosis Clinical Trials

Ekaterina V. Kurbatova^{1*}, Patrick P. J. Phillips^{2*}, Susan E. Dorman³, Erin E. Sizemore¹, Kia E. Bryant¹, Anne E. Purfield^{1,4}, Jessica Ricaldi¹, Nicole E. Brown¹, John L. Johnson^{5,6}, Carole L. Wallis⁷, Joseph P. Akol⁶, Oksana Ocheretina⁸, Nguyen Van Hung⁹, Harriet Mayanja-Kizza⁶, Madeleine Lourens¹⁰, Rodney Dawson¹¹, Nguyen Viet Nhung⁹, Samuel Pierre⁸, Yeukai Musodza¹², Justin Shenje¹³, Sharlaa Badal-Faesen¹⁴, Stalz Charles Vilbrun⁸, Ziyaad Waja¹⁵, Lakshmi Peddareddy¹, Nigel A. Scott¹, Yan Yuan¹, Stefan V. Goldberg^{1‡}, Susan Swindells¹⁶, Richard E. Chaisson¹⁷, Payam Nahid², and the AIDS Clinical Trials Group and the Tuberculosis Trials Consortium

¹U.S. Centers for Disease Control and Prevention, Atlanta, Georgia; ²UCSF Center for Tuberculosis, University of California San Francisco, San Francisco, California; ³Medical University of South Carolina, Charleston, South Carolina; ⁴United States Public Health Service Commissioned Corps, Rockville, Maryland; ⁵Case Western Reserve University, University Hospitals Cleveland Medical Center, Cleveland, Ohio; ⁶Uganda–Case Western Reserve University Research Collaboration, Kampala, Uganda; ⁷Lancet Laboratories and Bio Analytical Research Corporation South Africa (BARC SA), Johannesburg, South Africa; ⁸GHESKIO, Port-au-Prince, Haiti; ⁹Vietnam National Tuberculosis Program/National Lung Hospital, Hanoi, Vietnam; ¹⁰TASK, Cape Town, South Africa; ¹¹Division of Pulmonology, Department of Medicine, University of Cape Town and University of Cape Town Lung Institute, Cape Town, South Africa; ¹²University of Zimbabwe College of Health Sciences, Harare, Zimbabwe; ¹³South African Tuberculosis Vaccine Initiative, Cape Town, South Africa; ¹⁴Clinical HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa; ¹⁵Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa; ¹⁶University of Nebraska Medical Center, Omaha, Nebraska; and ¹⁷Johns Hopkins University School of Medicine, Baltimore, Maryland

ORCID IDs: 0000-0003-4636-3731 (E.V.K.); 0000-0002-6336-7024 (P.P.J.P.); 0000-0003-3232-0546 (E.E.S.); 0000-0003-0785-7800 (K.E.B.); 0000-0002-3184-0550 (N.E.B.); 0000-0003-1783-0476 (J.L.J.); 0000-0002-6736-3362 (O.O.); 0000-0003-3054-3930 (J.S.); 0000-0002-7306-5732 (S.B.-F.); 0000-0001-8717-9039 (Z.W.); 0000-0002-8513-7342 (S.V.G.); 0000-0001-5826-6037 (S.S.); 0000-0003-2811-1311 (P.N.).

Abstract

Rationale: We developed a standardized method, possible poor treatment response (PPTR), to help ascertain efficacy endpoints in Study S31/A5349 (NCT 02410772), an open-label trial comparing two 4-month rifampentine-based regimens with a standard 6-month regimen for the treatment of pulmonary tuberculosis (TB).

Objectives: We describe the use of the PPTR process and evaluate whether the goals of minimizing bias in efficacy endpoint assessment and attainment of relevant data to determine outcomes for all participants were achieved.

Methods: A PPTR event was defined as the occurrence of one or more prespecified triggers. Each PPTR required initiation of a standardized evaluation process that included obtaining multiple sputum samples for microbiology.

Measurements and Main Results: Among 2,343 participants with culture-confirmed drug-susceptible TB, 454 individuals

(19.4%) had a total of 534 individual PPTR events, of which 76.6% were microbiological (positive smear or culture at or after 17 wk). At least one PPTR event was experienced by 92.4% (133 of 144) of participants with TB-related unfavorable outcome and between 13.8% and 14.7% of participants with favorable and not-assessable outcomes. A total of 75% of participants with TB-related unfavorable outcomes had microbiological confirmation of failure to achieve a disease-free cure.

Conclusions: Standardized methodologies, such as our PPTR approach, could facilitate unbiased efficacy outcome determinations, improve discrimination between outcomes that are related and unrelated to regimen efficacy, and enhance the ability to conduct pooled analyses of contemporary trials.

Keywords: multicenter randomized trial; noninferiority; outcomes; TB; endpoints ascertainment bias

(Received in original form June 13, 2022; accepted in final form February 15, 2023)

*These authors contributed equally to this work.

‡Retired.

Supported by U.S. Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Division of Tuberculosis Elimination contracts 200-2009-32582, 200-2009-32593, 200-2009-32594, 200-2009-32589, 200-2009-32597, 200-2009-32598, 75D30119C06702, 75D30119C06701, 75D30119C06703, 75D30119C06222, 75D30119C06225, and 75D30119C06010; and by the National Institute of Allergy and Infectious Diseases of the NIH under award numbers UM1 AI068634, UM1 AI068636, and UM1 AI106701. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC, the National Institute of Allergy and Infectious Diseases, or the U.S. Department of Health and Human Services.

Am J Respir Crit Care Med Vol 207, Iss 10, pp 1376–1382, May 15, 2023

Copyright © 2023 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.202206-1118OC on February 15, 2023

Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the

Subject: In noninferiority tuberculosis (TB) trials, particular attention to quality, including ascertainment of outcomes, is critical. When adequate objective, often microbiological, data on treatment response are not available, efficacy outcome determinations are vulnerable to subjective clinical decisions that may lead to treatment changes and thereby affect efficacy endpoints. This can result in an observed dilution of the difference between treatments, which in a noninferiority trial may inflate the chance of a false-positive finding.

What This Study Adds to the

Field: We developed a standardized method, possible poor treatment response (PPTR), to help ascertain efficacy endpoints in an open-label noninferiority phase 3 trial of TB treatment, Study S31/A5349 (NCT 02410772). The PPTR process evaluates all available findings, regardless of their potential effect on study outcome, in order to answer most comprehensively a study's main question. The implementation of a standardized and widely accepted approach for the ascertainment of efficacy endpoints in late-phase TB treatment trials would optimize trial rigor and enhance the ability to conduct pooled analyses of contemporary trials.

A major goal of tuberculosis (TB) drug development is to identify regimens that can cure the disease with a shorter treatment duration without a loss in effectiveness.

Noninferiority designs are therefore commonly used in pivotal TB trials (1–3). Particular attention to quality, including ascertainment of outcomes, is critical when planning and conducting a noninferiority trial (2). When adequate objective data (often microbiological) on treatment response are not available, efficacy outcome determinations are vulnerable to subjective clinical decisions that may lead to treatment changes and thereby affect efficacy endpoints. This can result in an observed dilution of the difference between treatment that is protective in a superiority trial (reducing the chance of a false-positive result) but not necessarily in a noninferiority trial, where the chance of a false-positive result can become inflated (4). Blinding is a commonly used strategy to prevent knowledge of treatment allocation affecting the management of study participants or endpoint assessments. In therapeutic drug trials, blinding is typically accomplished through the use of inactive placebos intended to be indistinguishable from the investigational agent(s). Although blinding can mitigate bias, incorporation of placebos into a trial introduces challenges that must be weighed against the merits of their use.

In a recent phase 3 noninferiority multicenter trial evaluating novel TB treatment regimens, S31/A5349 (NCT 02410772), we faced a situation in which the use of placebos would have significantly increased daily pill burden, likely affecting participant acceptability and adherence, that did not outweigh potential benefits (see further discussion in References 3 and 5). Therefore, we designed a trial (3, 5) that did not incorporate placebos but instead incorporated a process of standardized data collection in the event of a possible poor treatment response (PPTR), intended to minimize bias in the efficacy endpoint. This process used prespecified criteria that were applied consistently and uniformly across all study arms.

The objectives of this analysis were the following: 1) to describe the use of the PPTR process during the conduct of Study S31/A5349 and identify which triggers were most important for identifying an unfavorable outcome for recommendations in future trials; and 2) to evaluate whether the goals of minimizing bias in efficacy endpoint assessment and attainment of relevant data to determine outcomes for all participants were achieved.

Methods

Study Design

Study S31/A5349 was an international, multicenter, randomized, controlled, open-label, three-arm, phase 3 noninferiority trial conducted at 34 sites in 13 countries (3, 5). The trial compared two 4-month (17 wk) rifapentine-based regimens with a standard 6-month (26 wk) regimen consisting of rifampin, isoniazid, pyrazinamide, and ethambutol (control) for the treatment of pulmonary TB. In one 4-month regimen, rifampin was replaced with rifapentine. In the other 4-month regimen, rifampin was replaced with rifapentine, and ethambutol was replaced with moxifloxacin, which was continued throughout treatment. The primary efficacy endpoint was TB disease-free survival at 12 months after study treatment assignment (3, 5). For each participant, a primary outcome status of favorable, unfavorable, or not assessable was assigned (3). Participants' outcomes, including TB-related unfavorable outcomes, were classified strictly according to the protocol and statistical analysis plan (SAP) definitions: a TB-related unfavorable outcome was defined as either 1) two consecutive positive cultures at or after Week 17 (all TB-related deaths met this definition); 2) not seen at Month 12, last culture positive; or 3) clinical diagnosis of TB recurrence and treatment restarted.

Author Contributions: Study conception and design: E.V.K., P.P.J.P., S.E.D., E.E.S., S.V.G., S.S., R.E.C., and P.N. Data collection: E.V.K., E.E.S., K.E.B., A.E.P., J.R., N.E.B., J.L.J., C.L.W., J.P.A., O.O., N.V.H., H.M.-K., M.L., R.D., N.V.N., S.P., Y.M., J.S., S.B.-F., S.C.V., Z.W., L.P., N.A.S., Y.Y., S.V.G., and P.N. Data analysis: P.P.J.P. and K.E.B. Data interpretation: E.V.K., P.P.J.P., S.E.D., S.S., R.E.C., and P.N. Drafting of the initial manuscript: E.V.K., P.P.J.P., S.E.D., and P.N. Critical review of the final draft of the manuscript: E.V.K., P.P.J.P., S.E.D., E.E.S., K.E.B., A.E.P., J.R., N.E.B., J.L.J., C.L.W., J.P.A., O.O., N.V.H., H.M.-K., M.L., R.D., N.V.N., S.P., Y.M., J.S., S.B.-F., S.C.V., Z.W., L.P., N.A.S., Y.Y., S.V.G., S.S., R.E.C., and P.N. Access and verification of underlying data: P.P.J.P., N.A.S., and K.E.B.

Correspondence and requests for reprints should be addressed to Payam Nahid, M.D., M.P.H., UCSF Center for Tuberculosis, University of California San Francisco, 1001 Potrero Avenue, 5K1, San Francisco, CA 94110. E-mail: pnahid@ucsf.edu.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Table 1. Possible Poor Treatment Response Triggers and Evaluation Procedures

Standardized PPTR Triggers	Standardized PPTR Evaluation Procedures
PPTR triggers driven by microbiological and clinical data Culture of sputum obtained at or after Week 17 is positive for <i>Mycobacterium tuberculosis</i> Smear microscopy of sputum obtained at or after Week 17 is positive for acid-fast bacilli Worsening signs or symptoms compatible with TB at or after Week 17 Radiographic worsening compatible with TB at or after Week 17 Treatment change PPTR triggers Site investigator is considering an extension of TB treatment beyond that of the participant's assigned regimen Site investigator is considering reinitiating any TB treatment after the participant has completed assigned study treatment Site investigator is considering a change in treatment for efficacy reasons (this does not apply to changes in treatment due to pregnancy or drug toxicity or to temporary drug rechallenge)	<ol style="list-style-type: none"> 1. Review of participant personal contact information 2. Review of adverse events and interval medical history 3. Symptom assessment 4. Concomitant medication assessment 5. Measurement of weight 6. Chest radiograph 7. At least THREE (3) SPUTA should be collected within 1 week. At least TWO of these sputa should be collected before changing or restarting TB treatment, and at least 4 h apart. At least ONE sputum should be a first morning specimen, if feasible. These sputa should be sent to the study laboratory for smear and culture. If <i>M. tuberculosis</i> is isolated in culture, drug susceptibility testing should be performed on one of the isolates. Then the isolate from each positive culture should be stored frozen 8. For participants consenting to collection of specimens for identification of potential biomarkers of TB treatment response: sputum, urine, and blood specimens should be obtained 9. Contact the central study clinician when PPTR evaluation is initiated and ideally before changing or restarting TB treatment* 10. Complete the PPTR case report form

Definition of abbreviations: PPTR = possible poor treatment response; TB = tuberculosis.

If a participant experienced any of the PPTR triggers, then all of the PPTR evaluation procedures should have been implemented regardless of treatment assignment.

*The central study clinician was blinded to participant treatment assignment. The purpose of this communication was to review the PPTR procedures and ensure adherence with them.

Table 2. Possible Poor Treatment Response Triggers and Evaluations for Possible Poor Treatment Response Events, Microbiologically Eligible Population (N=2,343)

	Total (N = 2,343)	Control (n = 768)	Rifapentine-Moxifloxacin Regimen (n = 791)	Rifapentine Regimen (n = 784)
Total PPTR events	534	139	179	216
Number of PPTR events per participant, median (range)		0 (0–3)	0 (0–4)	0 (0–4)
Among those with favorable outcome		0 (0–3)	0 (0–4)	0 (0–4)
Among those with unfavorable outcome		0 (0–3)	1 (0–3)	1 (0–3)
Among those with not-assessable outcome		0 (0–2)	0 (0–2)	0 (0–1)
Time to first PPTR event (10th centile, wk)	36.1	43.9	34.3	32.9
PPTR triggers*				
Smear positive at or after Week 17	213 (39.9)	43 (30.9)	86 (48.0)	84 (38.9)
Culture positive at or after Week 17	196 (36.7)	66 (47.5)	62 (34.6)	68 (31.5)
Worsening TB signs or symptoms at or after Week 17	140 (26.2)	33 (23.7)	39 (21.8)	68 (31.5)
Radiographic worsening at or after Week 17	34 (6.4)	8 (5.8)	7 (3.9)	19 (8.8)
Clinician considering permanent change in treatment	25 (4.7)	3 (2.2)	9 (5.0)	13 (6.0)
Other	18 (3.4)	7 (5.0)	5 (2.8)	6 (2.8)
Evaluations occurring within 1 wk of date of PPTR event				
No sputum sample	44 (8.2)	14 (10.1)	11 (6.1)	19 (8.8)
1 sputum sample	50 (9.4)	17 (12.2)	16 (8.9)	17 (7.9)
2 sputum samples	55 (10.3)	16 (11.5)	15 (8.4)	24 (11.1)
At least 3 sputum samples	385 (72.1)	92 (66.2)	137 (76.5)	156 (72.2)
Chest radiograph obtained after PPTR	442 (82.8)	120 (86.3)	148 (82.7)	174 (80.6)

For definition of abbreviations, see Table 1.

Table shows n (% of total PPTR events) unless otherwise specified.

*More than one trigger could be selected by the site on case report form.

Table 3. Summary of Possible Poor Treatment Response Events by Treatment Arm, Microbiologically Eligible Analysis Population (N = 2,343)

Total Number of PPTR Events	Participants with ≥1 PPTR Event; n (%) / Total in Population					
	All Participants	Participants with Favorable Outcome*	Participants with Unfavorable Outcome†	Participants with Unfavorable Outcome: TB-related‡	Participants with Unfavorable Outcome: Not TB-related§	Participants with Not-Assessable Outcome
Total	454 (19.4)/2,343	289 (14.7)/1,969	150 (56.6)/265	133 (92.4)/144¶	17 (14.0)/121	15 (13.8)/109
Control arm	119 (15.5)/768	90 (13.7)/656	24 (34.3)/70	18 (75.0)/24	6 (13.0)/46	5 (11.9)/42
RPT-MOX arm	151 (19.1)/791	101 (15.1)/668	46 (52.3)/88	43 (95.6)/45	3 (7.0)/43	4 (11.4)/35
RPT arm	184 (23.5)/784	98 (15.2)/645	80 (74.8)/107	72 (96.0)/75	8 (25.0)/32	6 (18.8)/32
Differences between arms, chi-square test	P < 0.001	P < 0.001	P = 0.004	P = 0.004	P = 0.096	P = 0.571

Definition of abbreviations: MOX = moxifloxacin; PPTR = possible poor treatment response; RPT = rifapentine; TB = tuberculosis.

*A favorable outcome corresponds with TB disease-free survival at 12 months after randomization. Favorable required that a participant did not meet criteria for unfavorable or not assessable, and either 1) had their latest sputum cultures at Month 12 negative for *Mycobacterium tuberculosis*; or 2) was without signs and symptoms of active TB at Month 12, and either unable to produce sputum or produced sputum that was contaminated without evidence of *M. tuberculosis*.

†A participant's outcome was classified as unfavorable if 1) cultures from two sputum specimens were positive for *M. tuberculosis* at or after Week 17 without an intervening negative culture; or 2) they died or were withdrawn or lost to follow-up during treatment; or 3) had a culture positive for *M. tuberculosis* when last seen; or 4) died from TB during post-treatment follow-up; or 5) received additional treatment for TB.

‡A TB-related unfavorable outcome was defined as either 1) two consecutive positive cultures at or after Week 17 (all TB-related deaths met this definition); 2) not seen at Month 12, last culture positive; or 3) clinical diagnosis of TB recurrence and treatment restarted.

§Not TB-related unfavorable outcome was defined as either 1) consent withdrawn during treatment; no adverse event reported; 2) treatment changed due to adverse event; 3) death during treatment (not met definition of TB-related death); 4) lost to follow-up during treatment; 5) consent withdrawn during treatment, after occurrence of adverse event; or 6) treatment changed or restarted for other reasons.

||A participant's outcome was classified as not assessable if they were not already classified as unfavorable and in addition 1) did not attend the Month 12 visit but were culture negative when last seen; or 2) had treatment changed because of pregnancy; or 3) died during follow-up with cause unrelated to TB; or 4) received additional treatment for TB after exogenous reinfection demonstrated by whole-genome sequencing; or 5) died from a violent or accidental death during treatment.

¶PPTR evaluation was not conducted for 11 (7.6%) of 144 participants with a TB-related unfavorable outcome; all of these participants were not seen at Month 12 but had a positive culture when last seen. A PPTR evaluation occurred for all other TB-related unfavorable outcomes.

The trial was approved by the CDC Institutional Review Board (IRB). Each participating institution provided for the review and approval of this protocol and its informed consent documents by a local IRB or ethics committee or relied formally on the CDC IRB approval.

Trial participants were monitored regularly according to a prespecified study schedule that included collection of sputum samples at 13 time points over the 18 months of study participation. All participants were weighed and questioned about new or worsened signs and symptoms of TB at all scheduled and unscheduled visits. All participants had a chest radiograph at the beginning and the end of study treatment and whenever clinically indicated. The microbiologically eligible analysis population included participants who, at study entry, had a culture positive for *Mycobacterium tuberculosis* that was not resistant to isoniazid, rifampin, or fluoroquinolones and were not randomized in violation of eligibility criteria. All analyses presented here used the microbiologically eligible analysis population unless otherwise stated.

PPTR Procedures

The foundational principle of the PPTR process was that any of 1) a set of prespecified laboratory and clinical findings; or 2) the site investigator's intention to change treatment constituted a "PPTR trigger" and required a prompt standardized evaluation, regardless of assigned treatment regimen and clinical impression of the local site study team. A PPTR event was defined as the presence of one or more of seven triggers requiring initiation of the evaluation process (Table 1). Participants could have more than one PPTR event during the trial. If a change in treatment was considered by the site investigator to be warranted, then PPTR procedures were to be performed before the treatment change, provided that the participant's clinical condition permitted the assessment. Investigators made the final determination of change of treatment and were not obliged to wait for results from PPTR evaluations. Participants whose treatment was changed continued to be followed in the study unless they withdrew consent.

PPTR Implementation

Staff at all study sites were trained on the PPTR triggers and evaluation procedures before the start of participant enrollment, and refresher trainings were provided at least

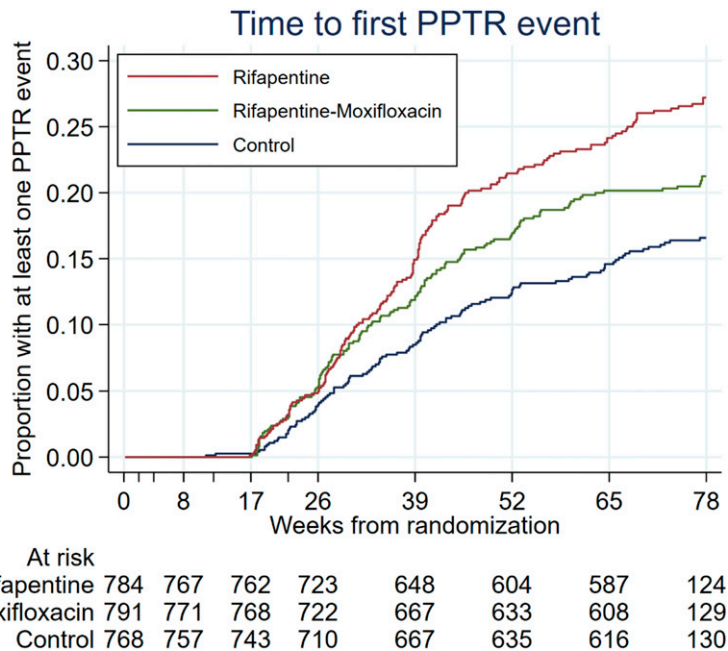


Figure 1. Timing of possible poor treatment response (PPTR) events. Kaplan-Meier curves show time from randomization to first PPTR event in the microbiologically eligible analysis population ($N = 2,343$).

annually. Each PPTR event was reported by the sites on an electronic case report form. The central study clinician at the data center reviewed each PPTR event within 1 business day from report and communicated with the sites to confirm PPTR procedure adherence or ask questions. The central study clinician was blinded to participant treatment assignment during review and communication with the site staff. Adherence to the PPTR process was reinforced and validated by a quality assurance activity that automatically discovered participants meeting criteria for a microbiological trigger but for whom a PPTR form had not been submitted (6). In such instances, the data center staff instructed site staff to verify the reported mycobacteriology result and to promptly conduct the PPTR evaluation if it was warranted per the protocol.

Statistical Methods

All analyses in this report used the primary microbiologically eligible analysis population (3), only excluding participants lacking culture confirmation of drug-susceptible TB. Chi-square tests were used for comparisons of proportions. Outcome determination was performed according to the SAP, previously published online as part of the primary paper supplementary

material protocol, with determination of TB- and non-TB-related outcomes as defined in the SAP (3). We plotted receiver operating characteristic curves for various combinations of PPTR triggers to predict TB-related unfavorable outcomes to visualize the effect on the balance between sensitivity and specificity. These analyses were based on the primary July 27, 2020 data extract that was used for the primary paper (3).

Results

PPTR Triggers

Among 2,343 participants included in the microbiologically eligible analysis population, 454 (19.4%) had a total of 534 individual PPTR events. PPTR events were most frequently triggered by positive sputum smear (213 events, 39.9%) or positive culture (196 events, 36.7%) at or after Week 17, or by the clinical trigger of worsening TB signs and symptoms at or after Week 17 (140 of events, 26.2%) (Table 2). Overall, site investigator consideration of a permanent treatment change was an uncommon trigger, occurring for 4.7% (25 of 534) of events overall. For each of the PPTR triggers, the sensitivity and specificity for TB-related unfavorable outcomes are shown in Figure E1 in the

online supplement. The area under the receiver operating characteristic curve increased from less than 0.7 for a single trigger, to 0.82 when smear and culture were considered (either or both positive considered positive), and to 0.88 when smear, culture, and radiographic findings were considered (Figure E1). At PPTR events that occurred among participants having a favorable or not-assessable outcome, the most common triggers were the same as overall: a positive smear (126 [37.1%] participants with a favorable outcome and 10 [58.8%] participants with not-assessable outcome) and a positive culture (126 [37.1%] participants with a favorable outcome and 6 [35.3%] participants with not-assessable outcome). Among 534 PPTR events, a total of 25 (4.7%) treatment change PPTR triggers were noted; among these 25, change in treatment or restart of treatment occurred in 18 (72%).

Adherence to Implementation of the PPTR Evaluation

Among 534 PPTR evaluations, the protocol-mandated minimum of three sputum specimens was obtained for 385 (72.1%) overall, including 66.2% (92 of 139) in the control group, 76.5% (137 of 179) in the rifapentine-moxifloxacin group, and 72.2% (156 of 216) in the rifapentine group. A chest radiograph was obtained for 82.8% (442 of 534) of events overall, including 86.3% (120 of 139) in the control group, 82.7% (148 of 179) in the rifapentine-moxifloxacin group, and 80.6% (174 of 216) in the rifapentine group.

Participants with PPTR Events and Outcomes Classification

The proportions of participants who had at least one PPTR event differed across study arms ($P < 0.001$) (Table 3); at least one PPTR event was experienced by 15.5% of participants in the control arm, 19.1% of participants in the rifapentine-moxifloxacin arm, and 23.5% of participants in the rifapentine arm. The time to first PPTR event was shortest in the rifapentine group and longest in the control group ($P < 0.001$; log rank test) (Table 1 and Figure 1). Among participants with a TB-related unfavorable outcome, at least one PPTR event was experienced by 92.4% (133 of 144) overall, including 75.0% (18 of 24) in the control group, 95.6% (43 of 45) in the rifapentine-moxifloxacin group, and 96.0% (72 of 75) in the rifapentine group ($P = 0.004$). Among

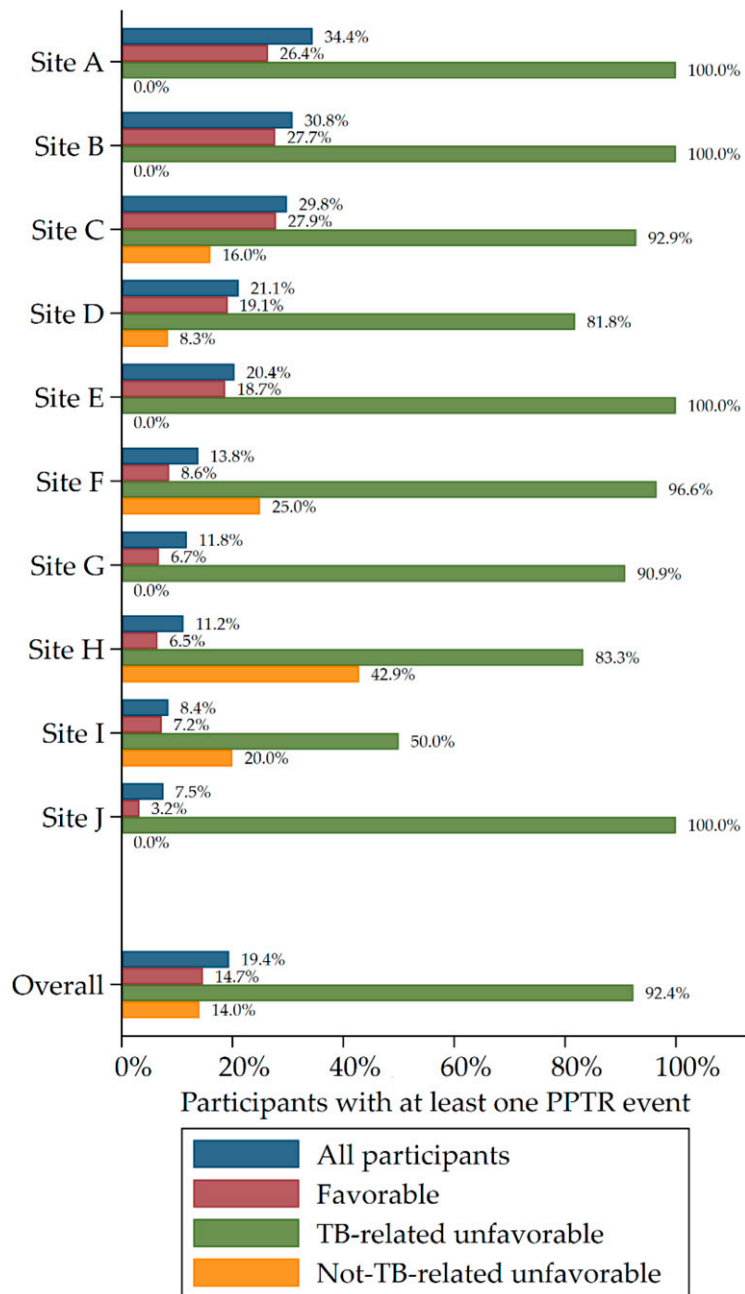


Figure 2. Summary of proportion of all participants with at least one PPTR event by site (including sites that enrolled at least 100 participants) in the microbiologically eligible analysis population. PPTR = possible poor treatment response; TB = tuberculosis.

the 144 participants with a TB-related unfavorable outcome, 12.5% (18 of 144) had a clinical diagnosis of TB recurrence resulting in treatment change without any microbiological confirmation, including 8.3% (2 of 24) in the control group, 17.8% (8 of 45) in the rifapentine-moxifloxacin group, and 10.7% (8 of 75) in the rifapentine group; all of these individuals had at least one PPTR evaluation performed. Among

the 144 participants classified per the protocol as having a TB-related unfavorable outcome, 11 (7.6%) participants (6 in the control arm, 2 in the rifapentine-moxifloxacin arm, and 3 in the rifapentine arm) did not have PPTR evaluation done; these 11 participants were not seen at Month 12 but had a positive culture when last seen before Month 12 (i.e., they were lost to follow-up and therefore could not be

found for PPTR evaluations). Among participants with a favorable outcome, at least one PPTR event was experienced by 14.7% (289 of 1,969) of participants overall, including 13.7% (90 of 656) in the control arm, 15.1% (101 of 668) in the rifapentine-moxifloxacin arm, and 15.2% (98 of 645) in the rifapentine group ($P < 0.001$) (Table 2). The number of participants with a PPTR event did differ by study site, ranging from 7.5% to 34.4% among sites that enrolled at least 100 participants (Figure 2). Among these 10 sites, the Spearman's rank correlation between proportion of TB-related unfavorable outcomes and proportion of participants with a PPTR event was 0.614 ($P = 0.059$).

Discussion

We describe the successful implementation of a standardized approach for the collection of objective data to support better primary outcome determination for late-phase multicenter TB trials. Several metrics point to high fidelity with the PPTR procedures during the conduct of the phase 3 trial. There were only 11 participants who did not undergo PPTR evaluation yet were classified as TB-related unfavorable, and all of these participants were lost to follow-up, such that PPTR evaluations were not possible. There were no participants classified as unfavorable owing to treatment change or retreatment in whom a PPTR evaluation was not done. Three sputum samples were collected within 1 week at 72.1% of PPTR events, and a chest radiograph was obtained after 82.8% of events.

Several key findings emerged from this analysis of the PPTR procedure implemented as a protocol-specified component of the S31/A5349 phase 3, open-label, noninferiority trial. There were relatively few TB-related unfavorable outcomes that were not microbiologically confirmed, not having positive cultures on two separate visits (36 of 144, 25%), thereby minimizing the frequency of more subjective outcomes and the likely impact of ascertainment bias in this open-label trial and strengthening the evidence for the finding of noninferiority. In another large phase 3 TB treatment trial (REMoxTB, a study for the "Rapid Evaluation of Moxifloxacin in the treatment of sputum smear positive tuberculosis") conducted in a similar patient population but without procedures analogous to the PPTR

process, 37% (78 of 209) of participants classified as TB-related unfavorable did not have microbiological confirmation (7). A total of 534 PPTR events occurred in 454 participants in our trial, in which only 144 of 2,343 participants experienced a TB-related unfavorable outcome, illustrating the additional work involved by sites to support the PPTR process. The occurrence of PPTR events varied by study site, likely reflecting the variability in occurrence of TB-related unfavorable outcomes that was also seen across study sites because of the differences in the severity of TB disease, including cavitary disease and initial sputum bacillary burden, and other participant characteristics (modest correlation, Spearman's rank correlation of 0.614).

Few PPTR evaluations were triggered by site investigator consideration of regimen change; however, multiple options were also selected on the case report form, so investigators may have omitted "considering changing treatment" if they had already noted that there were positive smears or cultures. In any case, we cannot conclude that this set of triggers should be omitted from the PPTR process in an open-label TB clinical trial, as it allows some insight into potential contribution of ascertainment bias to overall study findings. It is notable that the time to first PPTR event was shorter on experimental arms and, among participants with a favorable outcome, more PPTR events occurred on the experimental arms (15.2% and 15.1%) than on the control arm (13.7%), which likely reflects concerns of site

investigators, given they were not blinded to treatment allocation.

There are limitations to our findings regarding impact of the PPTR process on ascertainment bias in an open-label study. Most importantly, we cannot establish a cause-and-effect relationship because the PPTR process was applied to all participants and all treatment groups. Our experience was that blinding of the central study clinician was successful for distinguishing investigational arms but sometimes less so for control versus investigational arms (because of different treatment duration). Furthermore, not all participants completed all specified PPTR evaluations. We observed that the rank ordering of the arms in number of PPTR events analysis was the same as that for the overall number of unfavorable outcomes. This is not unexpected, because a goal of the PPTR process was to ensure completeness of data on which efficacy outcomes were based. However, this relationship underscores the potential importance of safeguards that prevent inadvertent or intentional use of PPTR time-to-event data as a means to glean insight into overall trial results.

In summary, the implementation of a standardized and widely accepted approach for the ascertainment of efficacy endpoints in late-phase TB treatment trials would improve quality that is so important in noninferiority trials, and this standardization will enhance the ability to conduct pooled analyses of contemporary trials. The PPTR process evaluates all

available findings, regardless of their potential effect on study outcome, in order to answer most comprehensively a study's main question without fear or favor. Standardized methodologies, such as our PPTR approach, could facilitate unbiased efficacy outcome determinations and improve discrimination between outcomes that are related and unrelated to regimen efficacy. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the study participants who contributed their time to this trial and site and local tuberculosis (TB) program staff who assisted in the clinical management of some study participants. They also thank Lara Hosey, clinical trials specialist, Richard Hafner, the Division of AIDS (DAIDS) medical officer, Kristine Coughlin, data manager, Akbar Shahkolahi, international program specialist, and Christopher Lane, laboratory technologist for support. The authors thank Phil LoBue, Carla Winston, and Jonathan Mermin for continued support of the TB Trials Consortium within the CDC; and Westat, Inc., and PPD, Inc. for on-site monitoring. Andrew Vernon, then working at CDC, provided conceptual and implementation support. Sanofi (Paris, France, and Bridgewater, New Jersey) donated rifapentine and all other study drugs, supported shipping of study drugs to all sites, and provided funding support for pharmacokinetic testing and preparation of the final Clinical Study Report in this collaborative study. The Division of Tuberculosis Elimination at the US Centers for Disease Control and Prevention served as the data and coordinating center for the study.

References

- Lienhardt C, Nunn A, Chaisson R, Vernon AA, Zignol M, Nahid P, *et al.* Advances in clinical trial design: weaving tomorrow's TB treatments. *PLoS Med* 2020;17:e1003059.
- U.S. Department of Health and Human Services; Food and Drug Administration; Center for Drug Evaluation and Research (CDER); Center for Biologics Evaluation and Research (CBER). Non-inferiority clinical trials to establish effectiveness: guidance for industry. 2016 [accessed 2022 May 6]. Available from: "<https://www.fda.gov/media/78504/download>" Non-Inferiority Clinical Trials to Establish Effectiveness Guidance for Industry (fda.gov).
- Dorman SE, Nahid P, Kurbatova EV, Phillips PPJ, Bryant K, Dooley KE, *et al.*; AIDS Clinical Trials Group, Tuberculosis Trials Consortium. Four-month rifapentine regimens with or without moxifloxacin for tuberculosis. *N Engl J Med* 2021;384:1705–1718.
- Phillips PPJ, Glidden DV. Noninferiority trials. In: Piantadosi S, Meinert CL, editors. Principles and practice of clinical trials. Cham: Springer; 2021. pp 1297–1323.
- Dorman SE, Nahid P, Kurbatova EV, Goldberg SV, Bozeman L, Burman WJ, *et al.*; AIDS Clinical Trials Group and the Tuberculosis Trials Consortium. High-dose rifapentine with or without moxifloxacin for shortening treatment of pulmonary tuberculosis: study protocol for TBTC study 31/ACTG A5349 phase 3 clinical trial. *Contemp Clin Trials* 2020;90:105938.
- Bryant KE, Yuan Y, Engle M, Kurbatova EV, Allen-Blige C, Batra K, *et al.*; AIDS Clinical Trials Group, Tuberculosis Trials Consortium. Central monitoring in a randomized, open-label, controlled phase 3 clinical trial for a treatment-shortening regimen for pulmonary tuberculosis. *Contemp Clin Trials* 2021;104:106355.
- Gillespie SH, Crook AM, McHugh TD, Mendel CM, Meredith SK, Murray SR, *et al.*; REMoxTB Consortium. Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. *N Engl J Med* 2014;371:1577–1587.