ORIGINAL ARTICLE

Effectiveness of Bedaquiline Use beyond Six Months in Patients with Multidrug-Resistant Tuberculosis

Letizia Trevisi[1], Miguel A. Hernán[2], Carole D. Mitnick[1,3,4], Uzma Khan[5], Kwonjune J. Seung[1,3,4], Michael L. Rich[1,3,4], Mathieu Bastard[6], Helena Huerga[6], Nara Melikyan[6], Sidney A. Atwood[3], Zaza Avaliani[7], Felix Llanos[8,9], Mohammad Manzur-ul-Alam[10], Khin Zarli[11], Amsalu Bekele Binegdie[12], Sana Adnan[13], Arusyak Melikyan[14], Alain Gelin[15], Afshan K. Isani[16], Dmitry Vetushko[17], Zhenisgul Daugarina[18], Patrick Nkundanyirazo[19], Fauziah Asnely Putri[5], Charles Vilbrun[20], Munira Khan[21], Catherine Hewison[22]*, Palwasha Y. Khan[23]*, and Molly F. Franke[1]*; on behalf of the endTB Study Team

[1]Department of Global Health and Social Medicine, Harvard Medical School, Boston, Massachusetts; [2]CAUSALab, Departments of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts; [3]Division of Global Health Equity, Brigham and Women's Hospital, Boston, Massachusetts; [4]Partners in Health, Boston, Massachusetts; [5]Interactive Research and Development Global, Singapore, Singapore; [6]Field Epidemiology Department, Epicentre, Paris, France; [7]National Center for Tuberculosis and Lung Diseases, Tbilisi, Georgia; [8]Unidad de Tuberculosis, Hospital Nacional Dos de Mayo, Lima, Peru; [9]Instituto de Investigaciones en Ciencias Biomedicas, Universidad Ricardo Palma, Lima, Peru; [10]Interactive Research and Development, Dhaka, Bangladesh; [11]Médecins sans Frontières, Yangon, Myanmar; [12]Department of Internal Medicine, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia; [13]Indus Hospital and Health Network, Karachi, Pakistan; [14]Médecins sans Frontières, Yerevan, Armenia; [15]Zanmi Lasante, Port-au-Prince, Haiti; [16]National Core Research Group, Stop TB Partnership, Islamabad, Pakistan; [17]Republican Scientific and Practical Centre of Pulmonology and Tuberculosis, Minsk, Belarus; [18]Astana City Center of Phthisiopulmonology, Astana, Kazakhstan; [19]Partners in Health, Maseru, Lesotho; [20]GHESKIO, Port-au-Prince, Haiti; [21]Interactive Research and Development, Durban, South Africa; [22]Medical Department, Médecins sans Frontières, Paris, France; and [23]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

ORCID ID: 0000-0002-4890-5728 (M.F.F.).

Abstract

Rationale: Current recommendations for the treatment of rifampicin- and multidrug-resistant tuberculosis include bedaquiline (BDQ) used for 6 months or longer. Evidence is needed to inform the optimal duration of BDQ.

Objectives: We emulated a target trial to estimate the effect of three BDQ duration treatment strategies (6, 7–11, and ⩾12 mo) on the probability of successful treatment among patients receiving a longer individualized regimen for multidrug-resistant tuberculosis.

Methods: To estimate the probability of successful treatment, we implemented a three-step approach comprising cloning, censoring, and inverse probability weighting.

Measurements and Main Results: The 1,468 eligible individuals received a median of 4 (interquartile range, 4–5) likely effective drugs. In 87.1% and 77.7% of participants, this included linezolid and clofazimine, respectively. The adjusted probability of successful treatment was 0.85 (95% confidence interval [CI], 0.81–0.88) for 6 months of BDQ, 0.77 (95% CI, 0.73–0.81) for 7–11 months, and 0.86 (95% CI, 0.83–0.88) for ⩾12 months. Compared with 6 months of BDQ, the ratio of treatment success was 0.91 (95% CI, 0.85–0.96) for 7–11 months and 1.01 (95% CI, 0.96–1.06) for ⩾12 months. Naive analyses that did not account for bias revealed a higher probability of successful treatment with ⩾12 months (ratio, 1.09 [95% CI, 1.05–1.14]).

Conclusions: BDQ use beyond 6 months did not increase the probability of successful treatment among patients receiving longer regimens that commonly included new and repurposed drugs. When not properly accounted for, immortal person–time bias can influence estimates of the effects of treatment duration. Future analyses should explore the effect of treatment duration of BDQ and other drugs in subgroups with advanced disease and/or receiving less potent regimens.

Keywords: rifampicin-resistant TB; target trial; duration; inverse probability weighting; endTB observational study

(Received in original form November 18, 2022; accepted in final form February 17, 2023)

*These authors contributed equally to this work.

Correspondence and requests for reprints should be addressed to Molly F. Franke, Sc. D., Department of Global Health and Social Medicine, Harvard Medical School, 641 Huntington Avenue, Boston, MA 02115. E-mail: molly_franke@hms.harvard.edu.

Am J Respir Crit Care Med Vol 207, Iss 11, pp 1525–1532, Jun 1, 2023

Originally Published in Press as DOI: 10.1164/rccm.202211-2125OC on February 21, 2023
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: The World Health Organization recommends that bedaquiline (BDQ) be used in longer (18–20 months) individualized regimens for 24 weeks, the duration studied in two randomized trials. Although evidence supports the safety of BDQ use beyond 24 weeks, there is no evidence on the effectiveness of BDQ beyond this period.

What This Study Adds to the Field: We found no evidence that BDQ use beyond 6 months increased the probability of successful treatment among patients receiving longer regimens that commonly included new and repurposed drugs, such as BDQ, delamanid, linezolid, and clofazimine. Although future analyses should examine the effectiveness of longer durations of BDQ among patient subgroups at high risk of unfavorable outcomes, these results suggest that in regimens with enough potent companion drugs, 6 months of BDQ may be sufficient for many patients. To address this question, we used the target trial approach, an analytic framework novel to the study of tuberculosis treatments. This framework enabled the articulation of a clear, relevant research question and the design of an analysis that corrected for potential biases common to studies of treatment duration. Results from naive analyses using prevailing methods showed bias. The target trial framework may be a useful tool for future analyses of tuberculosis treatment cohorts.

Since its approval nearly a decade ago, the anti–tuberculosis (TB) drug bedaquiline (BDQ) has contributed importantly to safer, more tolerable, and more effective treatment for rifampicin-resistant (RR) and multidrug-resistant (MDR) TB. In 2018, the World Health Organization (WHO) recommended that BDQ be used in longer (18–24 mo) individualized regimens for 24 weeks, the duration studied in two pivotal randomized trials (1, 2), and cited a lack of evidence on the safety and effectiveness of BDQ beyond this period (3). Two years later, in 2020, WHO determined that BDQ use beyond 24 weeks was safe. However, although identifying the optimal duration of BDQ had been a key research priority, no new guidance was provided for BDQ duration. This was due to limited data on the effectiveness of use beyond 24 weeks and potential residual confounding by indication in an analysis of observational data (4). In other words, sicker patients and/or those with weaker regimens were more likely to receive BDQ for longer durations; inadequately accounting for this confounding could introduce a bias that would attenuate or invert a beneficial effect of prolonged use.

Current recommendations for the treatment of RR/MDR-TB include regimens with durations longer than 6 months; therefore, there is a critical need for high-quality evidence to inform the optimal duration of BDQ. This evidence should be generated from analyses that fully avail of detailed longitudinal data sets and analytic methods that address the time-varying confounding by indication (5, 6) that precluded conclusions on the efficacy of BDQ use beyond 6 months in the 2020 WHO review. Also critical are methods that account for the potential immortal time bias that arises in analyses of treatment duration because people who survive for longer can be treated for longer (7). Though commonly used in other fields (8), appropriate methods to adjust for time-varying confounding (e.g., inverse probability [IP] weighting) and immortal person–time bias have rarely been applied in analyses of TB treatment cohorts (5).

In this study, we examined the comparative effectiveness of prolonged BDQ use (i.e., for 7–11 or ⩾12 mo relative to 6 mo), on favorable end-of-treatment outcomes. To do so, we emulated a target trial (9) to account for key biases and compared our findings with those of analyses that did not adjust for these biases.

Methods

We emulated a (hypothetical) target trial using data from the endTB (Expand new drug markets for TB) observational study, a multicountry cohort of 2,789 patients with RR/MDR-TB who were treated with a longer (18–24 mo) regimen, containing BDQ and/or delamanid. Patients enrolled in the endTB observational study were treated according to WHO and national guidelines, under routine programmatic conditions, with additional guidance from the endTB clinical guide (10). Enrollment took place in 17 countries between 2015 and 2018, and patients were followed through the end of treatment. A common protocol was used across sites (11). Drug prescription start and stop dates, and reasons for regimen changes were tracked. Clinically relevant adverse events were reported, adverse events were graded by the reporting physician according to the Médecins sans Frontières severity scale, and serious adverse events were reported to the Médecins sans Frontières pharmacovigilance unit, which routinely reconciled data with the electronic medical record at each site (12). For this analysis, we excluded individuals from the Democratic People's Republic of North Korea because of differences in diagnosis and treatment

Supported by Unitaid and by National Institute of Allergy and Infectious Diseases grant R01AI46095 (L.T., M.L.R., K.J.S., S.A.A., M.A.H., C.D.M., and M.F.F.). The funders had no role in the conceptualization, analysis or presentation of findings of this study. Bedaquiline donations made from Janssen to the Global Drug Facility were used for patients in the endTB observational study. Donations of delamanid from Otsuka were used for initial patients enrolled in the endTB observational study. The companies from which drug donations were received did not have any role in study design, data analyses, data interpretation, or manuscript writing.

Author Contributions: Study design: K.J.S., H.H., U.K., M.L.R., C.D.M., C.H., M.B., P.Y.K., and M.F.F. Analysis design: M.A.H., M.F.F., and L.T. Study implementation: Z.A., F.L., M.M.-u.-A., K.Z., A.B.B., S.A., N.M., A.M., A.G., A.K.I., D.V., Z.D., P.N., F.A.P., C.V., and M.K. Data management: L.T. and S.A.A. Statistical analyses: L.T. and M.F.F. Data interpretation: L.T., M.F.F., M.A.H., C.D.M., M.L.R., K.J.S., U.K., P.Y.K., and C.H. First draft: L.T. and M.F.F. Critical review of manuscript: all authors.

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. Protocol and Emulation of a Target Trial of Bedaquiline Duration Using Data from a Longitudinal Cohort of Patients Undergoing Treatment for Rifampicin-Resistant/Multidrug-Resistant Tuberculosis

Characteristics	Target Trial Protocol	Emulated Trial Using Observational Data
Eligibility	• Confirmed RR/MDR-TB • Initiated BDQ within the first month of MDR treatment • Completed 24 wk of BDQ within the first 27 wk of treatment • BDQ was likely to be effective at the end of Week 23 • Not treated in the Democratic People's Republic of Korea	Same
Treatment strategies	1. BDQ for 6 mo 2. BDQ for 7–11 mo 3. BDQ for ⩾12 mo • The content and duration of background regimen (i.e., other TB drugs included in the multidrug regimen) are determined on the basis of clinical judgment. • Participants can deviate from their assigned treatment arm (BDQ duration) for <2 wk for any prescribed reason (e.g., adverse event, out of stock). • Adverse events resulting in suspension of BDQ for ⩾2 wk are managed according to clinical judgment. • BDQ can be reinitiated after stopping when clinically indicated.	Same
Treatment assignment	Patients are randomly assigned to one of the three strategies and are aware of the strategy to which they are assigned.	Patients are assigned (via clones) to all the strategies compatible with their data at time zero
Follow-up	Follow-up starts in Week 24 of BDQ and ends at the end of treatment.	Same
Outcome	TB treatment success (i.e., cure or treatment completion). Those who die, who are lost from treatment, and in whom treatment fails are considered to have unsuccessful outcomes.	Same
Causal contrast	Intention-to-treat effect, per-protocol effect	Observational analogue of the per-protocol effect
Statistical analysis	Intention-to-treat analysis: probabilities of the outcome under each assigned strategy are compared via ratios and differences. Per-protocol analysis is the same, but patients are censored when they deviate from their assigned strategy, and potential selection bias is adjusted for using IP weighting.	Same per-protocol analysis, except that patient clones are used

Definition of abbreviations: BDQ = bedaquiline; IP = inverse probability; MDR = multidrug-resistant; RR = rifampicin-resistant; TB = tuberculosis.

(including the use of shortened regimens) compared with the rest of the cohort.

Specification of the Target Trial

The target trial (9) would enroll patients aged 10–78 years with RR/MDR-TB who initiated BDQ within a month of treatment initiation, who completed 24 weeks of BDQ doses within the first 27 weeks of treatment, and in whom there was no prior exposure or evidence of resistance to BDQ, at Week 23 (Table 1). At Week 24, each eligible participant would be randomized to one of three durations of BDQ: 6 months (from Week 24 through Week 28), 7–11 months (from Week 29 through Week 47), or ⩾12 months (from Week 48 through the end of treatment). Under all three strategies, BDQ could be stopped after an adverse event, with subsequent management (e.g., drug substitutions or additions) left to the clinician's discretion. BDQ could be reinitiated after the stoppage indicated by the assigned strategy, if clinically directed. Clinician-directed BDQ interruptions of less than 14 days, for any reason (e.g., drug out of stock) would not be considered protocol deviations. Background regimens would follow WHO guidelines and local norms and would be adjusted if clinically indicated. Participants would be followed from assignment until the end of treatment, at which time treatment outcomes would be calculated on the basis of 2013 WHO guidance (13, 14). Successful treatment, the outcome of interest, would include cure and treatment completion, whereas unsuccessful treatment would include outcomes of death, loss to follow-up, and treatment failure.

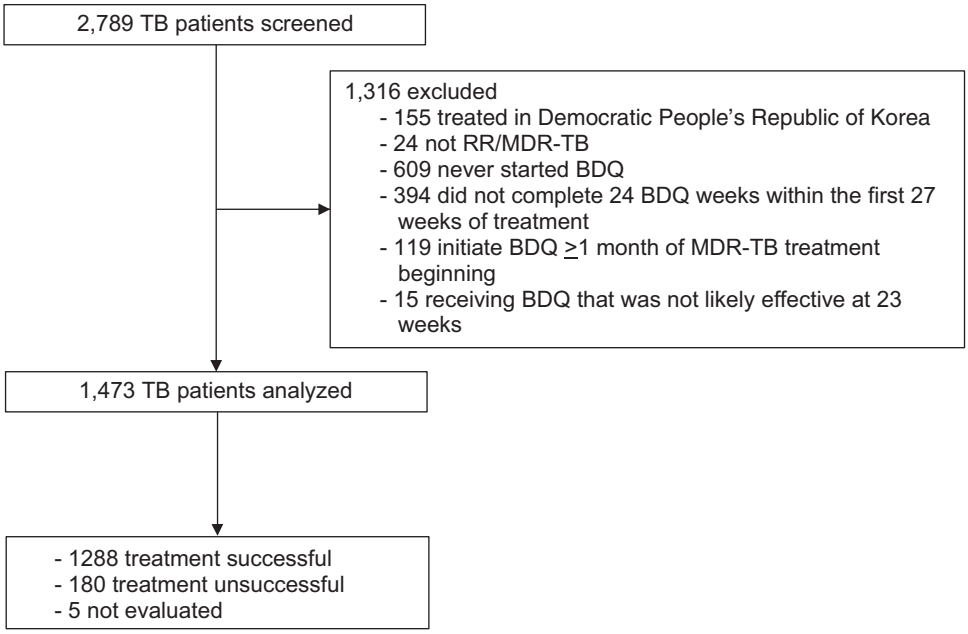

Figure 1. Flowchart of analysis inclusion, endTB (Expand New Drug Markets for TB) Cohort, 2015–2018. BDQ = bedaquiline; MDR = multidrug-resistant; RR = rifampicin-resistant; TB = tuberculosis.

Emulation of the Target Trial

We emulated the target trial using the endTB observational data set (Table 1). We used the same eligibility criteria as the target trial to define the patient population. In total, 1,473 participants in the endTB study met these criteria (Figure 1), of whom 5 were eventually censored from the analysis because of outcomes recorded as "not evaluated." Time zero of follow-up was the 24th week of BDQ prescription. We then followed a three-step analysis procedure to compare the three treatment strategies (9, 15). First, we created a data set with three copies (clones) of each patient. Each copy was assigned to one of the three BDQ treatment strategies. This cloning avoids the immortal time bias that can occur if the treatment strategy is assigned on the basis of the observed duration of treatment. Second, we artificially censored each patient copy if/when it deviated from its assigned strategy. For example, clones assigned to the strategy BDQ for 7–11 months were censored if they stopped BDQ for the first time before Month 7 or after Month 11 for a reason other than an adverse event. Third, we used IP weighting to adjust for potential selection bias introduced by the artificial censoring (16). Informally censored individuals transfer their weight to similar (in terms of specified covariates) uncensored individuals to create a hypothetical population in which everybody followed their assigned strategy. Figure E1 in the online supplement provides an overview of the cloning, censoring, and IP weighting steps, including how they were implemented and which potential biases they address.

Statistical Analysis

IP-weighted model. We fit an IP-weighted logistic regression model for the probability of treatment success among uncensored patient copies. The logistic model included age (continuous), sex, and the most recent values of the following covariates assessed at Week 23 (i.e., baseline): cumulative adherence $>$ 80% (binary, with missing counted as $\leqslant$80%), extensive disease (i.e., cavitary disease with a sputum smear result of 2+ or 3+), number of severe adverse events (binary: none vs. one or more), functional status (17) (categorical: fully active, restricted in physically strenuous activity but ambulatory, ambulatory with full self-care vs. limited self-care, or completely disabled), body mass index (BMI) $<$ 18.5 kg/m^2, culture result (positive or negative), number of drugs in the regimen (count variable), prescription of cycloserine, and prescription of the following drugs that were likely to be effective in the patient: fluoroquinolone, linezolid, clofazimine, and a second-line injectable. A drug was considered likely effective if *1*) all reported testing (phenotypic or genotypic) to that drug confirmed susceptibility or *2*) no resistance to the drug was reported and the patient had not previously received the drug for 1 month or more. Otherwise, the drug was not considered likely effective. We use this model to estimate the unconditional probability of successful treatment. These probabilities were compared via ratios and differences. Ninety-five percent confidence intervals were computed using the nonparametric bootstrap with 500 samples.

Calculation of IP weights. To estimate the denominator of the IP weights, we fit a logistic model for the probability of being uncensored (i.e., remaining on BDQ) for each patient–week with week of follow-up (linear and quadratic terms), the baseline covariates listed above, and time-varying covariates (*see* Table E1). Time-varying covariates were updated when values changed and included the cumulative number of prior severe adverse events, current use of cycloserine, current use of a second-line injectable that was likely to be effective, BMI $<$ 18.5 kg/m^2, culture result, and functional status. We estimated stabilized IP weights with a numerator estimated using a second logistic model for probability of being uncensored with baseline covariates only. The distributions of the weights are shown in Table E2. Missing data were rare (5–10% for most key variables).

Table 2. Characteristics of 1,468 Patients Receiving Bedaquiline for at Least 24 Weeks, endTB (Expand New Drug Markets for TB) Cohort, 2015–2018

Characteristic	*n* (%)*	Missing [*n* (%)]
Assessed at treatment initiation		
Age at treatment initiation, yr, median (IQR; range)	34 (26–45; 10–78)	—
Female sex	531 (36.2)	—
HIV infection	95 (6.5)	1 (0.1)
Hepatitis C	139 (9.5)	3 (0.2)
Previous TB treatment with second-line drugs	1,181 (80.5)	—
Resistance profile		
RR/MDR-TB with FQ and injectable both sensitive	203 (13.8)	—
RR/MDR-TB without testing to FQ and injectable	77 (5.2)	—
RR/MDR-TB with injectable resistance, FQ sensitive	182 (12.4)	—
RR/MDR-TB with FQ resistance, injectable sensitive	402 (27.4)	—
RR/MDR-TB with FQ and injectable both resistant	554 (37.7)	—
Not tested for RR/MDR-TB	30 (2.0)	—
Assessed at 23 wk		
Complete data on adherence	1,294 (88.2)	—
Adherence rate > 80% among those with complete adherence (*n* = 1,294)	1,163 (89.9)	—
Positive sputum culture	78 (5.3)	13 (0.9)
Extensive disease	—	149 (10.1)
No cavitary disease, no smear 3+	460 (31.3)	—
Cavitary disease, smear <3+	845 (57.6)	—
No cavitary disease, smear 3+	4 (0.3)	—
Cavitary disease, smear 3+	10 (0.7)	—
BMI < 18.5 kg/m^2	448 (30.5)	7 (0.5)
Functional status	—	124 (8.4)
Fully active	726 (49.5)	—
Ambulatory	468 (31.9)	—
Capable of self-care	112 (7.6)	—
Limited self-care	27 (1.8)	—
Completely disabled	11 (0.7)	—
FQ	724 (49.3)	—
Likely effective†	494 (33.7)	—
Linezolid	1,323 (90.1)	—
Likely effective	1,278 (87.1)	—
Clofazimine	1,209 (82.4)	—
Likely effective	1,140 (77.7)	—
Cycloserine	831 (56.6)	—
Likely effective	266 (18.1)	—
Delamanid	311 (21.2)	—
Likely effective	309 (21.1)	—
Number of likely effective drugs, median (IQR; range)	4 (4–5; 1–9)	—

Definition of abbreviations: BMI = body mass index; FQ = fluoroquinolone; IQR = interquartile range; MDR = multidrug-resistant; RR = rifampicin-resistant; TB = tuberculosis.
*Unless otherwise noted.
†A drug was considered likely effective if *1*) all reported testing (phenotypic or genotypic) of that drug confirmed susceptibility or *2*) no resistance to the drug was reported and the patient had not previously received the drug for 1 month or more. Otherwise, the drug was not considered likely effective.

Unless otherwise specified, we used missing indicator variables to account for missingness in baseline and time-varying covariates (18).

To examine the impact of increasing covariate adjustment, we report estimates from unadjusted analyses, age- and sex-adjusted analyses, baseline-adjusted analyses, and baseline-adjusted and IP-weighted analyses. For comparison purposes, we reanalyzed data using methods that did not account for immortal time bias by categorizing patients according to their observed BDQ treatment duration (6, 7–11, and ⩾12 mo). All analyses were conducted using SAS 9.4 (SAS Institute).

Research Ethics

The endTB observational study protocol was approved by central ethics review committees for each consortium partner, and local ethics approval was obtained in all endTB countries. Participants provided written informed consent for inclusion in the observational cohort.

Results

Baseline Characteristics

The 1,468 participants had a median age of 34 years (interquartile range, 26–45 yr) and were predominantly men (63.8%; Table 2). HIV and hepatitis C infection cooccurred with TB in 6.5% and 9.5% of patients, respectively. Only a small fraction of participants had extensive disease (0.7%). Almost a third of participants (30.5%) had BMIs < 18.5 kg/m^2, and more than 50% had some degree of impairment to their activities of daily living. The median number of likely effective drugs included in the regimen was 4 (interquartile range, 4–5), and 87.1%, 77.7%, and 33.7% of patients were receiving linezolid, clofazimine, and a fluoroquinolone that was likely to be effective, respectively (Table 2).

Primary Effectiveness Analysis

In both the baseline-adjusted analyses and IP-weighted analyses, the predicted probability of successful treatment was 0.85 (95% confidence interval [CI], 0.81–0.88) for 6 months of BDQ, 0.77 (95% CI, 0.73–0.81) for 7–11 months, and 0.86 (95% CI, 0.83–0.88) for ⩾12 months (Table 3). Compared with 6 months of BDQ, the treatment success ratio was 0.91 (95% CI, 0.85–0.96) for 7–11 months and 1.01 (95% CI, 0.96–1.06) for ⩾12 months (Table 4). Unadjusted and age- and sex-adjusted estimates were more extreme than adjusted estimates (Table 4). Treatment success differences were −0.08 (95% CI, −0.13 to −0.03) and 0.01 (95% CI, −0.03 to 0.05) for 7–11 and ⩾12 months, respectively (Table 4).

Table 3. Probabilities of End-of-Treatment Success under Several Bedaquiline Duration Strategies, endTB (Expand New Drug Markets for TB) Cohort, 2015–2018

	Unweighted Models			
	Unadjusted	Adjusted for Age and Sex	Baseline Adjusted*	Weighted and Baseline Adjusted*
BDQ Duration	Prob. of Treatment Success (95% CI)	Prob. of Treatment Success (95% CI)	Prob. of Treatment Success (95% CI)	Prob. of Treatment Success (95% CI)
6 mo	0.88 (0.85–0.90)	0.87 (0.85–0.90)	0.85 (0.81–0.88)	0.85 (0.81–0.88)
7–11 mo	0.75 (0.71–0.80)	0.75 (0.71–0.80)	0.77 (0.73–0.81)	0.77 (0.73–0.81)
⩾12 mo	0.84 (0.82–0.87)	0.84 (0.82–0.87)	0.86 (0.83–0.88)	0.86 (0.83–0.88)

Definition of abbreviations: BDQ = bedaquiline; CI = confidence interval; Prob = probability.
All analyses shown here were conducted on the cloned data set and were therefore adjusted for immortal time bias.
*The model was adjusted for age (years; continuous), sex, cumulative adherence > 80% (binary, with missing counted as ⩽80%), extensive disease (i.e., cavitary disease with a sputum smear result of 2+ or 3+), number of severe adverse events (binary: none vs. one or more), functional status (categorical: fully active, restricted in physically strenuous activity but ambulatory, ambulatory with full self-care vs. limited self-care, or completely disabled), body mass index < 18.5 kg/m^2, culture result (positive or negative), number of drugs in the regimen (count variable), prescription of cycloserine, and prescription of the following drugs that were likely to be effective in the patient: fluoroquinolone, linezolid, clofazimine, and a second-line injectable.

Naive Analysis

A baseline-adjusted analysis that introduced immortal person–time bias yielded a treatment success ratio of 1.09 (95% CI, 1.05–1.14) for ⩾12 months of BDQ compared with 6 months (Table 5).

Discussion

In patients receiving 18- to 20-month individualized regimens (with a median of four drugs), BDQ used for longer than 6 months did not improve the probability of treatment success over the 85% achieved with 6 months of treatment. The high proportion of success observed here confirmed—or improved on—results from several landmark trials and observational studies among patients treated with 6 months of BDQ in conventional regimens lasting 18–24 months (19–22). More recently, the TB-PRACTECAL (Pragmatic Clinical Trial for a More Effective Concise and Less Toxic MDR-TB Treatment Regimen[s]) and STREAM-2 (The Evaluation of a Standard Treatment Regimen of Anti-Tuberculosis Drugs for Patients with Multi-Drug-Resistant Tuberculosis) trials have revealed the efficacy of shorter regimens, lasting 6–9 months, that contain BDQ (23, 24). Taken together, these results suggest that in regimens with enough potent companion drugs, 6 months of BDQ may be sufficient for many patients. To date, it is unknown whether certain subsets of patients (i.e., those with cavitary and highly smear-positive disease [25, 26], those with fluoroquinolone resistance [27, 28], and those on less potent regimens) could benefit from a longer duration of BDQ or of other drugs. Addressing this question will likely require

Table 4. Estimated Effectiveness of Several Bedaquiline Duration Strategies on Successful End-of-Treatment Outcome, endTB (Expand New Drug Markets for TB) Cohort, 2015–2018

	Unweighted Models			
	Unadjusted	Adjusted for Age and Sex	Baseline Adjusted*	Weighted and Baseline Adjusted*
BDQ Duration	SR (95% CI)	SR (95% CI)	SR (95% CI)	SR (95% CI)
6 mo	Ref.	Ref.	Ref.	Ref.
7–11 mo	0.86 (0.80 to 0.91)	0.86 (0.81 to 0.91)	0.91 (0.85 to 0.96)	0.91 (0.85 to 0.96)
⩾12 mo	0.96 (0.92 to 1.01)	0.97 (0.92 to 1.01)	1.01 (0.96 to 1.06)	1.01 (0.96 to 1.06)
	SD (95% CI)	SD (95% CI)	SD (95% CI)	SD (95% CI)
6 mo	Ref.	Ref.	Ref.	Ref.
7–11 mo	−0.12 (−0.17 to −0.07)	−0.12 (−0.17 to −0.07)	−0.08 (−0.13 to −0.03)	−0.08 (−0.13 to −0.03)
⩾12 mo	−0.03 (−0.07 to 0.008)	−0.03 (−0.07 to 0.009)	0.01 (−0.03 to 0.05)	0.01 (−0.03 to 0.05)

Definition of abbreviations: BDQ = bedaquiline; CI = confidence interval; Ref = reference; SD = success difference; SR = success ratio.
All analyses shown here were conducted on the cloned data set and were therefore adjusted for immortal time bias.
*The model was adjusted for age (years; continuous), sex, cumulative adherence > 80% (binary, with missing counted as ⩽80%), extensive disease (i.e., cavitary disease with a sputum smear result of 2+ or 3+), number of severe adverse events (binary: none vs. one or more), functional status (categorical: fully active, restricted in physically strenuous activity but ambulatory, ambulatory with full self-care vs. limited self-care, or completely disabled), body mass index < 18.5 kg/m^2, culture result (positive or negative), number of drugs in the regimen (count variable), prescription of cycloserine, and prescription of the following drugs that were likely to be effective in the patient: fluoroquinolone, linezolid, clofazimine, and a second-line injectable.

Table 5. Naive Analysis: Estimated Effectiveness of Bedaquiline Duration on Successful End-of-Treatment Outcome, endTB (Expand New Drug Markets for TB) Cohort, 2015–2018

BDQ Duration	*n*	Successful Outcome [*n* (%)]	Adjusted SR (95% CI)
6 mo	538	470 (87.4)	Ref.
7–11 mo	272	215 (79)	0.93 (0.88–0.99)
⩾12 mo	658	603 (91.6)	1.09 (1.05–1.14)

Definition of abbreviations: BDQ = bedaquiline; CI = confidence interval; Ref = reference; SR = success ratio.

pooling across studies to obtain sufficient sample sizes of the subgroups of interest.

The research question addressed here, whether longer use of BDQ improves the probability of treatment success, was inspired by the 2019 WHO treatment guideline update, which listed a similar question as one of four priority areas guiding the review and analysis of evidence. Although our analysis focused on BDQ, the optimal durations of most other second-line TB drugs (e.g., clofazimine, delamanid, fluoroquinolones) have never been studied. It is therefore possible that replicated for other individual drugs in the context of BDQ-containing regimens, the analyses performed here would also find a lack of evidence supporting an overall benefit of longer exposure. The recommended administration of BDQ had been limited to 6 months, both because this was the duration at which it was tested in pivotal trials (1) and because of initial concerns about BDQ cardiotoxicity, which led to a black-box warning. However, subsequent prospective studies have shown that the risk of serious corrected QT interval prolongation, arrhythmia, and sudden death is quite low, even when used in combination with other corrected QT interval–prolonging anti-TB drugs (delamanid, fluoroquinolones, and clofazimine) (12, 29, 30). Other second-line drugs (such as linezolid, clofazimine, and second-line injectables for as short as 8 weeks) do have toxicity or tolerability issues (24). In circumstances in which treatment is longer than 6 months, BDQ prolongation may have a favorable benefit/risk profile compared with these and other drugs: its extension could enable reduced exposure to other more toxic drugs, while maintaining overall regimen effectiveness and improving the patient experience. Our findings should not be taken to mean that there is no role for BDQ use beyond 24 weeks. Determining the optimal durations of individual drugs (and regimens) to improve effectiveness and avert toxicity will require larger studies.

Specifically, large randomized trials with high adherence and no loss to follow-up would yield effect estimates that are not expected to be affected by confounding. In the absence of such trials, observational analyses that explicitly emulate trials estimate treatment effects under assumptions of no residual or unmeasured confounding, no selection or information bias, and correct model misspecification (31, 32). Our explicit target trial emulation allowed us to answer a clearly specified research question and to avoid immortal person–time bias. The impact of immortal time bias is evident when comparing the effect estimates between our analysis (success ratio, 1.01; 95% CI, 0.96–1.06) and the "naive" analysis (success ratio, 1.09; 95% CI, 1.05–1.14); the latter would have led to the erroneous conclusion that ⩾12 months of BDQ increased the probability of successful treatment. The potential for confounding is also evident: the unadjusted and age- and sex-adjusted effect estimates suggested a greater benefit than those adjusted for baseline covariates. Further adjustment for time-varying confounders in the IP-weighted analyses had little impact on the effect estimates.

We estimated a small decrease in the probability of successful treatment for 7–11 months compared with 6 months of BDQ (success difference, −0.08 [95% CI, −0.13 to −0.03]). It is unlikely that this is due to additional exposure to BDQ, considering its excellent efficacy and safety profile. Instead, this finding could be easily explained by random variability or misclassification of treatment adherence. In the programmatic setting, data on TB treatment adherence are typically reported by health workers via monthly treatment cards, which may overestimate adherence (33) and result in some missing values (12% in our cohort). Furthermore, the adherence data collected were for treatment overall, so we lacked data specific to BDQ. This misclassification of adherence would disproportionately bias estimates for the 7- to 11-month group because BDQ was often extended by the number of missed doses during the first 6 months. Those who missed three or more weeks of BDQ (consecutively or intermittently) during the first 6 months of treatment would have been classified in the 7- to 11-month group and may have experienced a lower probability of end-of-treatment success (34, 35) than those who completed their full courses of treatment in 6 months.

Conclusions

We did not find much evidence that treatment with BDQ beyond 28 weeks increased the probability of treatment success among patients receiving longer regimens that typically consisted of at least four likely effective new and repurposed drugs, such as clofazimine, linezolid, and delamanid. Future analyses should focus on subgroups with advanced disease, extensively drug-resistant strains, and/or less potent regimens that could potentially benefit from longer durations. This is among the first observational analyses of drug duration for RR/MDR-TB to appropriately account for time-varying confounding and immortal person–time bias. Collaborative efforts to develop and share longitudinal data sets will be critical to enabling rigorous analyses of observational data in the future. As longitudinal data sets of RR/MDR-TB become more available, the target trial framework will be a useful tool for specifying clear research questions and avoiding immortal time bias, with the overall goal of improving the evidence informing patient care and, ultimately, reducing TB-related morbidity and mortality. ■

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

Acknowledgment: The authors acknowledge and thank their funders, Unitaid, for financing the endTB project, including the observational study. The authors are indebted to the patients who participated in the endTB observational study and to the clinicians and program staff members of participating national TB programs. The authors thank endTB staff members at Partners in Health, Médecins sans Frontières, Epicentre, and Interactive Research and Development.

References

1. Diacon AH, Pym A, Grobusch MP, de los Rios JM, Gotuzzo E, Vasilyeva I, *et al.*; TMC207-C208 Study Group. Multidrug-resistant tuberculosis and culture conversion with bedaquiline. *N Engl J Med* 2014;371:723–732.
2. Pym AS, Diacon AH, Tang S-J, Conradie F, Danilovits M, Chuchottaworn C, *et al.*; TMC207-C209 Study Group. Bedaquiline in the treatment of multidrug- and extensively drug-resistant tuberculosis. *Eur Respir J* 2016;47:564–574.
3. World Health Organization. WHO treatment guidelines for rifampicin- and multidrug-resistant tuberculosis, 2018 update. Geneva, Switzerland: World Health Organization; 2018 [accessed 2023 Jan 3]. Available from: https://www.who.int/news-room/events/detail/2018/07/16/default-calendar/who-treatment-guidelines-for-rifampicin–and-multidrug-resistant-tuberculosis-2018-update.
4. World Health Organization. WHO consolidated guidelines on tuberculosis: module 1: prevention: tuberculosis preventive treatment. Geneva, Switzerland: World Health Organization; 2020 [accessed 2023 Jan 3]. Available from: https://www.who.int/publications/i/item/9789240001503.
5. Rodriguez CA, Sy KTL, Mitnick CD, Franke MF. Time-dependent confounding in tuberculosis treatment outcome analyses: a review of a source of bias. *Am J Respir Crit Care Med* 2020;202:1311–1314.
6. Franke MF, Mitnick CD. Time for a change: considering regimen changes in analyses of observational drug-resistant TB treatment cohort data. *Int J Tuberc Lung Dis* 2020;24:1151–1155.
7. Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ* 2018;360:k182.
8. Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding—a systematic review of the literature. *Int J Epidemiol* 2019;48:254–265.
9. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016;183:758–764.
10. Partners in Health, Médecins sans Frontières, Interactive Research and Development. endTB clinical guide. 2018 [accessed 2023 Jan 3]. Available from: https://www.endtb.org/.
11. Khan U, Huerga H, Khan AJ, Mitnick CD, Hewison C, Varaine F, *et al.* The endTB observational study protocol: treatment of MDR-TB with bedaquiline or delamanid containing regimens. *BMC Infect Dis* 2019;19:733.
12. Hewison C, Khan U, Bastard M, Lachenal N, Coutisson S, Osso E, *et al.* Safety of treatment regimens containing bedaquiline and delamanid in the endTB cohort. *Clin Infect Dis* 2022;75:1006–1013.
13. Zeng C, Mitnick C, Hewsion C, Bastard M, Khan P, Seung KJ, *et al.* Concordance of three approaches for operationalizing outcome definitions for multidrug-resistant tuberculosis. *Int J Tuberc Lung Dis* 2023;27:34–40.
14. World Health Organization. Definitions and reporting framework for tuberculosis—2013 revision: updated December 2014 and January 2020. Geneva, Switzerland: World Health Organization; 2013 [accessed 2023 Jan 3]. Available from: http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf.
15. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016;79:70–75.
16. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168:656–664.
17. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, *et al.* Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;5:649–655.
18. Groenwold RHH, White IR, Donders ART, Carpenter JR, Altman DG, Moons KGM. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. *CMAJ* 2012;184:1265–1269.
19. Conradie F, Diacon AH, Ngubane N, Howell P, Everitt D, Crook AM, *et al.*; Nix-TB Trial Team. Treatment of highly drug-resistant pulmonary tuberculosis. *N Engl J Med* 2020;382:893–902.
20. Ahmad N, Ahuja SD, Akkerman OW, Alffenaar JC, Anderson LF, Baghaei P, *et al.*; Collaborative Group for the Meta-Analysis of Individual Patient Data in MDR-TB Treatment–2017. Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet* 2018;392:821–834.
21. Borisov SE, Dheda K, Enwerem M, Romero Leyet R, D'Ambrosio L, Centis R, *et al.* Effectiveness and safety of bedaquiline-containing regimens in the treatment of MDR- and XDR-TB: a multicentre study. *Eur Respir J* 2017;49:1700387.
22. Ndjeka N, Schnippel K, Master I, Meintjes G, Maartens G, Romero R, *et al.* High treatment success rate for multidrug-resistant and extensively drug-resistant tuberculosis using a bedaquiline-containing treatment regimen. *Eur Respir J* 2018;52:1801528.
23. Nyang'wa BT, Berry C, Kazounis E, Motta I, Parpieva N, Tigay Z, *et al.*; TB-PRACTECAL Study Collaborators. A 24-week all-oral regimens for rifampicin resistant tuberculosis. *N Engl J Med* 2022;387:2331–2343.
24. Goodall RL, Meredith SK, Nunn AJ, Bayissa A, Bhatnagar AK, Bronson G, *et al.*; STREAM Study Collaborators. 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;400:1858–1868.
25. Imperial MZ, Nahid P, Phillips PPJ, Davies GR, Fielding K, Hanna D, *et al.* A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nat Med* 2018;24:1708–1715.
26. Franke MF, Khan P, Hewison C, Khan U, Huerga H, Seung KJ, *et al.* Culture conversion in patients treated with bedaquiline and/or delamanid: a prospective multicountry study. *Am J Respir Crit Care Med* 2021;203:111–119.
27. Mitnick CD, Shin SS, Seung KJ, Rich ML, Atwood SS, Furin JJ, *et al.* Comprehensive treatment of extensively drug-resistant tuberculosis. *N Engl J Med* 2008;359:563–574.
28. Farhat MR, Jacobson KR, Franke MF, Kaur D, Murray M, Mitnick CD. Fluoroquinolone resistance mutation detection is equivalent to culture-based drug sensitivity testing for predicting multidrug-resistant tuberculosis treatment outcome: a retrospective cohort study. *Clin Infect Dis* 2017;65:1364–1370.
29. Brust JCM, Gandhi NR, Wasserman S, Maartens G, Omar SV, Ismail NA, *et al.* Effectiveness and cardiac safety of bedaquiline-based therapy for drug-resistant tuberculosis: a prospective cohort study. *Clin Infect Dis* 2021;73:2083–2092.
30. Dooley KE, Rosenkranz SL, Conradie F, Moran L, Hafner R, von Groote-Bidlingmaier F, *et al.*; AIDS Clinical Trials Group (ACTG) A5343 DELIBERATE Study Team. QT effects of bedaquiline, delamanid, or both in patients with rifampicin-resistant tuberculosis: a phase 2, open-label, randomised, controlled trial. *Lancet Infect Dis* 2021;21:975–983.
31. Rodriguez CA, Lodi S, Horsburgh CR, Bastard M, Hewison C, Huerga H, *et al.* Selection bias in multidrug-resistant tuberculosis cohort studies assessing sputum culture conversion. *PLoS One* 2022;17:e0276457.
32. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA* 2022;328:2446–2447.
33. Prabhu A, Agarwal U, Tripathy JP, Singla N, Sagili K, Thekkur P, *et al.* "99DOTS" techno-supervision for tuberculosis treatment—a boon or a bane? Exploring challenges in its implementation at a tertiary centre in Delhi, India. *Indian J Tuberc* 2020;67:46–53.
34. Bastard M, Sanchez-Padilla E, Hewison C, Hayrapetyan A, Khurkhumal S, Varaine F, *et al.* Effects of treatment interruption patterns on treatment success among patients with multidrug-resistant tuberculosis in Armenia and Abkhazia. *J Infect Dis* 2015;211:1607–1615.
35. Cavanaugh JS, Kazennyy BY, Nguyen ML, Kiryanova EV, Vitek E, Khorosheva TM, *et al.* Outcomes and follow-up of patients treated for multidrug-resistant tuberculosis in Orel, Russia, 2002–2005. *Int J Tuberc Lung Dis* 2012;16:1069–1074.