



Totality of outcomes: A different paradigm in assessing interventions for treatment of tuberculosis



Grace Montepiedra^{a,1,*}, Courtney M. Yuen^{b,1}, Michael L. Rich^b, Scott R. Evans^a

^a Center for Biostatistics in AIDS Research and Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA

^b Division of Global Health Equity, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA

ARTICLE INFO

Article history:

Received 30 December 2015

Revised 5 May 2016

Accepted 8 May 2016

Keywords:

Tuberculosis

Treatment outcome

Risk–benefit assessment

ABSTRACT

Conventional analytic methods used for tuberculosis (TB) outcomes research use standardized outcomes definitions and assess safety and efficacy separately. These methods are subject to important limitations. Conventionally utilized outcome definitions fail to capture important aspects of patients' treatment experience and obscure meaningful differences between patients. Assessing safety and efficacy separately fails to yield an objective risk–benefit comparison to guide clinical practice. We propose to address these issues through an analytic approach based on prioritized outcomes. This approach enables a more comprehensive and integrated assessment of TB interventions. It simultaneously considers a “totality of outcomes”, including clinical benefit, adverse events, and quality of life. These composite outcomes are ranked terms of overall desirability and compared using statistical methods for ordinal outcomes. Here we discuss the application of this approach to TB research, the considerations involved with prioritizing TB treatment outcomes, and the statistical methods involved in comparing prioritized outcomes.

© 2016 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Studies assessing different treatment strategies for tuberculosis (TB) typically use binary outcomes (e.g., successful versus unsuccessful treatment, death versus survival) based on a standardized set of outcome definitions that were established to report TB program data to the World Health Organization. Five mutually exclusive outcomes are defined: cured, treatment completed, treatment failed, died, and lost to follow-up [1] (Table 1).¹ Treatment success is typically defined as either cured or treatment completed, which may poorly reflect how well a treatment works and how it contributes to patient well-being. Moreover, these classifications are subject to several limitations when used for TB treatment research, as they obscure meaningful differences between individual patient outcomes.

Firstly, these definitions do not consider side effects during treatment, so patients who complete treatment without any major side effects and patients who complete treatment but suffer irreparable hearing loss are equivalently classified. Secondly, the definitions do not consider the condition of a patient at the end of the observation period. Patients who complete treatment are classified as *treatment successes* even if they are faring poorly clinically with worsening radiographic findings at the end of treatment. Thirdly, the definitions do not capture risk of relapse, which occurs after the end of the prescribed treatment period but is arguably integral to the definition of cure. Fourthly, for patients who are classified as failing treatment, the definitions do not capture the possibility for retreatment. Because the first event that occurs is used to define the treatment outcome [2], patients in the “treatment failed” category may include those who ultimately died during the observation period and those who were ultimately cured [3]. And finally, because the definitions only describe patients' status at a single endpoint, they are ill-suited for incorporating indicators related to the treatment experience, such as the length of treatment, the pill burden, the dosing schedule, or the mode of administration (i.e., injectable versus oral).

An alternative analytic approach that could address the challenge of differentiating patient outcomes based on all meaningful

* Corresponding author at: Francois Xavier Bagnoud (FXB) Building Room 517, 651 Huntington Avenue, Boston, MA 02115-6017, USA. Tel.: +1 617 432 1141; fax: +1 617 432 3163.

E-mail address: gmontepie@sdac.harvard.edu (G. Montepiedra).

¹ These authors contributed equally to the work.

¹ This list excludes the categories of “Not evaluated,” which means that no treatment outcome has been assigned, and “Treatment success,” which is the sum of “Cured” and “Treatment completed.”

Table 1
World Health Organization reporting definitions for tuberculosis (TB) treatment outcomes.

Outcome	Definition for patients treated for TB susceptible to rifampin	Definition for patients treated for TB resistant to at least rifampin (including multidrug-resistant TB)
Cured	A pulmonary TB patient with bacteriologically confirmed TB at the beginning of treatment who was smear- or culture-negative in the last month of treatment and on at least one previous occasion	Treatment completed as recommended by the national policy without evidence of failure AND three or more consecutive cultures taken at least 30 days apart are negative after the intensive phase of treatment
Treatment completed	A TB patient who completed treatment without evidence of failure BUT with no record to show that sputum smear or culture results in the last month of treatment and on at least on previous occasion were negative, either because tests were not done or because results are unavailable	Treatment completed as recommended by the national policy without evidence of failure BUT no record that three or more consecutive culture taken at least 30 days apart are negative after the intensive phase of treatment
Treatment failed	A TB patient whose sputum smear or culture is positive at month 5 or later during treatment	Treatment terminated or need for permanent regimen change of at least two antituberculosis drugs because of one of four reasons (fully described in guidelines), which are related to lack of bacteriologic response, development of additional drug resistance, or adverse reactions to drug(s)
Died	A patient who dies for any reason before starting or during the course of treatment	A patient who dies for any reason during the course of treatment
Lost to follow-up	A TB patient who did not start treatment or whose treatment was interrupted for 2 consecutive months or more	A patient whose treatment was interrupted for 2 consecutive months or more

Table adapted from World Health Organization, "Definitions and reporting framework for tuberculosis – 2013 revision" [1].

comparisons is one based on prioritized outcomes. Prioritized outcomes approaches consider each individual patient's treatment experience with respect to multiple types of clinical outcomes during the entire period of observation (i.e., a "totality of outcomes") and then rank patients according to their overall treatment experience. Formal statistical comparisons are used to compare groups of patients based on the ranks of their totality of outcomes. This idea was first proposed in the statistical literature by Chuang-Stein in the context of clinical trials of antihypertensive drugs [4]. Since then, a body of methodological work has been produced in different disease areas [5–14]. More recently, Evans et al described an adaptation of this approach in the context of antibiotic stewardship trials [15]. In this concept paper, we describe how prioritized outcome approaches can be used to assess a totality of outcomes for TB treatment.

2. Example 1: a prioritized outcomes approach to risk–benefit analysis of TB treatments

To illustrate the advantage of a prioritized outcome approach, we present an example using a highly simplified scheme for ranking outcomes of patients treated for multidrug-resistant (MDR) TB. Many of the drugs available for treatment of MDR-TB are known to have substantial toxicity, and MDR-TB treatment regimens are poorly tolerated by patients. Clinicians are forced to subjectively weight the risks and benefits of using a regimen that may offer a greater chance of cure but results in a higher risk of adverse events.

Let us consider two regimens, A and B, each used to treat 300 patients, and producing the simplified outcome distributions shown in Fig. 1. Regimen B is associated with a significantly higher treatment success rate compared to regimen A (73% versus 65%, relative risk [RR] for treatment success = 1.31, 95% confidence interval [CI] 1.03–1.67), but also a significantly higher prevalence of serious adverse events (50% versus 40%, RR for serious adverse events = 1.20, 95% CI 1.04–1.39). Thus, a comparison based purely on clinical benefit would favor Regimen B, while a comparison based purely on toxicity would favor Regimen A. The question arises: Does the clinical benefit derived from choosing Regimen B outweigh the higher risk of serious adverse events associated with it?

A prioritized outcomes approach allows comparison of both indicators simultaneously and, thereby, directly addresses this risk–benefit question. One must first rank the desirability of

Regimen A			Regimen B		
	Treatment success	No treatment success		Treatment success	No treatment success
No serious adverse event	85	95	No serious adverse event	140	10
Serious adverse event(s)	110	10	Serious adverse event(s)	80	70

Treatment success rate: 65%
Serious adverse event rate: 40%

Treatment success rate: 73%
Serious adverse event rate: 50%

Fig. 1. Hypothetical distribution of outcomes and serious adverse events among patients treated with two regimens.

patient outcomes. In this case, let us consider: Treatment success without adverse event > treatment success with adverse event > lack of treatment success without adverse event > lack of treatment success with adverse event. Categorizing the 300 patients in each group into these four categories, then comparing the ranks in the two groups using the Wilcoxon rank sum tests, favors Regimen B with a p-value of 0.018. The estimated probability that a randomly selected patient taking Regimen B will have a better score than a patient from Regimen A is 55.4% (95% confidence interval [CI]: 52.8–57.9%) when all pairwise comparisons are included in the estimation, with half a point added to the numerator of the estimate whenever a tie occurs.

Thus, while comparing clinical benefit and toxicity separately yields contradictory information about which regimen may be preferable, a prioritized outcome approach suggests that Regimen B may be better overall, given these outcome distributions.

3. Prioritizing outcomes for TB research

The example above presented a simplistic outcome ranking scheme for illustrative purposes, but in actuality, the outcome ranking scheme could be much more complex. Developing this ranking scheme is the first and most important step in applying a prioritized outcome approach. It is important to acknowledge from the outset that the act of ranking is inherently subjective and different aspects of the treatment experience may be more important to consider depending on the research question and study context. Therefore, it is critical to achieve consensus in creating this ranking scheme before proceeding with analysis.

A method that has been used to validate prioritized outcome rankings for HIV [8] and cardiovascular disease [5] is to use consensus ranking to inform development of rule-based ranking

schemes. First, a panel of expert clinicians (or clinicians and patients) is used to subjectively rank the outcomes of a set of patients. These results are then used to inform development of multiple possible rule-based ranking schemes, which apply hierarchical sets of rules to differentiate patients first based on primary outcome indicators, then use secondary outcome indicators to differentiate among patients with the same primary outcome ranking. The performance of each rule-based ranking scheme is compared against the experts' rankings using Spearman's rank correlation, and the set of rules that produces a ranking most similar to the clinicians' judgement is identified. An alternative approach would be to use a Delphi process for achieving consensus among experts, or among experts and patients [16,17].

To illustrate some of the complexities involved in creating a rule-based ranking scheme, we present one possible ranked list of primary clinical outcomes, which seeks to integrate treatment completion, bacteriologic response, and clinical response. In this list, 1 represents the most desirable outcome, and 6 the worst outcome.

- (1) Treatment completed with bacteriologic evidence of a sustained cure.
- (2) Treatment completed without bacteriologic evidence of sustained cure, but with radiologic improvement or resolution of symptoms.
- (3) Treatment completed without bacteriologic evidence of sustained cure, and with no radiologic improvement or resolution of symptoms
- (4) Treatment not completed, but patient did not die during the set observation period, and sufficient bacteriological evidence was available to determine that there was no relapse within the observation period.
- (5) Treatment not completed, and patient did not die while receiving treatment, but either the patient relapsed within the set observation period, or insufficient bacteriologic evidence was available to determine absence of relapse.
- (6) Death attributable to TB at any time during set observation period.

One major question that this list elicits is how to rank patients in Outcome categories 3 and 4 relative to one another. While treatment completion is generally seen as superior to lack of treatment completion, one could argue that patients with Outcome 4 fared better than patients experiencing Outcome 3, who completed a clinically ineffective treatment. If there is no clear distinction in terms of superiority or inferiority between adjacent categories, then they could be combined. In addition, once secondary outcomes such as side effects are taken into account, the situation might become more complex. As a secondary outcome measure, side effects would be expected to differentiate among patients experiencing the same primary outcome, but not to change the relative rankings of patients with different primary outcomes. However, differences of opinion could exist on whether a patient who experienced severe side effects but had bacteriologic evidence of cure fared better than a patient who experienced no side effects but lacked bacteriologic evidence of cure. The inherent subjectivity of the ranking process requires that any ranking scheme used for analysis first be validated, and ideally, sensitivity analyses would be conducted to determine the effect of re-ordering ranks among contested sets of outcomes.

In addition to side effects, other aspects of the treatment experience that may be incorporated into a ranked list of totality of outcomes include whether a patient is left with disease-associated disability at the end of treatment, whether a patient acquired additional drug resistance during treatment, the total length of treatment received, and the number of weeks the patient was unable to work or attend school. However, one has to be careful incorpo-

rating criteria that are of lesser importance into a ranking scheme, as large differences in lesser criteria may obscure a meaningful difference in a more important criterion, or even possibly result in a different regimen being favored [18].

4. Example 2: evaluation of TB program performance using prioritized outcomes

The following example illustrates the added value of analyses that consider a totality of outcomes. Consider two TB programs with 75% treatment success rates (Fig. 2). Program A lacks a robust system for microbiologic follow-up and drug susceptibility testing (DST). As a result, a substantial proportion of patients complete ineffective regimens, and many treatment failures are either lost to follow-up or die. Program B has a system of microbiologic follow-up and DST for those at risk of treatment failure, so many patients on a failing regimen are ultimately switched to effective second-line regimens.

Fig. 2 shows the hypothetical distribution of outcomes based on the WHO classification, as well as a new set of seven ranked totality of outcomes, which are used to compare Programs A and B using a prioritized outcomes approach. Using the WHO definitions, here is no significant difference in treatment success rates between the two strategies (75% for both, relative risk [RR] for treatment success = 1.00, 95% confidence interval [CI] 0.85–1.17). However, comparison of the ranks between the two groups using the Wilcoxon rank sum test favors Program B, with a p-value of 0.038. The estimated probability that a patient in Program B will

	Program A	Program B
Treatment success	75	75
Treatment failure	20	20
Died	5	5

Rank	Scenario	WHO outcome category	Program A	Program B
1	Treatment completed with bacteriologic evidence of cure	Cured	30	35
2	Treatment completed with clinical improvement but no bacteriologic evidence of cure	Treatment completed	25	30
3	Evidence of smear positivity at month 5, then DST results became available; patient switched to second-line regimen, and ultimately cured	Treatment failure	0	15
4	Treatment completed with worsening clinical signs, but no evidence of smear positivity at month 5 (e.g., no test performed)	Treatment completed	20	10
5	Evidence of smear positivity at month 5, but patient's regimen not changed; patient completed ineffective regimen.	Treatment failure	15	5
6	Evidence of smear positivity at month 5, but patient's regimen not changed; patient then died.	Treatment failure	5	0
7	Patient died before month 5 smear result was known.	Died	5	5

Fig. 2. Hypothetical distribution of outcomes among 100 patients in each of two TB programs with equivalent treatment success rates.

have a better rank than a patient in Program A is 58.4% (95% confidence interval [CI]: 55.5–61.6% when all pairwise comparisons are included in the estimation.

We see that reliance on the conventionally utilized outcome definitions may not reveal differences in what ultimately happens to patients, as illustrated in this hypothetical example that aims to determine whether there is added value in a strategy of monitoring patients to assess the need for regimen changes during TB treatment. Hence, a totality of outcomes approach that differentiates among these different possibilities can show a difference in program performance even in the absence of increased “treatment success”.

5. Overview of statistical methods

Prioritized outcomes analyses use standard methods for analyzing ordinal data, including Wilcoxon’s rank sum test or Cochran–Mantel–Haenszel’s chi-square. As illustrated in the examples above, the treatment effect can be quantified by estimating the probability that a randomly selected patient in one arm will get a higher score than a randomly selected patient in the other arm based on comparison of all possible patient pairs. Confidence intervals around these estimates can be obtained using bootstrapping methods. Alternative measures of treatment effect based on pairwise comparisons of patients between the two treatment arms have also been proposed more recently, such as Buyse’s [10] “proportion in favor” and Pocock et al.’s “win ratio” quantities [11].

Ordinal logistic regression can also be used to compare ranked outcomes [19]. A constant odds ratio for all cumulative levels of the ranked outcomes can be estimated, as long as the data satisfy the proportional odds assumption. Otherwise, a model with nonconstant odds ratios can be fit, although this approach would yield separate treatment effect estimates for the different cumulative outcome categories. An advantage of the regression approach is that one can readily adjust for the presence of potential confounders or significant covariates, which is particularly important when analyzing data from non-randomized studies or observational cohorts. Alternative approaches when the proportional odds model fits poorly are also discussed by Agresti [19].

For example, we use ordinal logistic regression to compare the two regimens described in Example 1 above. We find that the proportional odds assumption does not hold for these data (score test p -value < 0.0001). However, the constant odds ratio estimate from logistic regression modeling using cumulative logits is still useful in providing an overall treatment effect measure [19]. In this case, the odds ratio (OR) of 1.44 (95% CI: 1.07–1.92) shows that Regimen B is generally favored. Modeling with nonproportional odds shows that Regimen B has significantly higher odds of treatment success without a serious adverse event than Regimen A (OR = 2.21, 95% CI: 1.58–3.10) and higher odds of treatment success with or without a serious adverse event (OR = 1.48, 95% CI: 1.05–2.10); however, Regimen B is associated with significantly lower odds of avoiding the worst outcome of treatment failure with a serious adverse event (OR = 0.11, 95% CI: 0.06–0.23).

One issue likely to be encountered in TB outcomes research, given the lengthy duration of treatment, is how to compare outcomes for patients who did not remain under observation for the entire prescribed observation period. One method for dealing with the complications of a censored outcome is to compare patient pairs over the common follow-up period [9].

6. Discussion

Conventionally utilized methods of assessing TB treatment outcomes do not capture the multiple dimensions of the treatment

experience that are meaningful in determining an optimal treatment approach. Conventional analytic methods that use binary treatment outcomes or rely on the standardized WHO outcome categories when comparing treatment strategies can fail to detect meaningful differences between patient experiences and final outcomes. Determining optimal treatment strategies will require methods capable of simultaneously considering measurements of efficacy, safety, and patient quality of life. As a complementary approach to conventional methods, prioritizing and evaluating the totality of outcomes over the long course of a patient’s treatment and follow-up could help provide a more comprehensive comparison of different treatment approaches.

In the applications discussed, we presented two examples to provide a clear picture of the methodology and its application to TB treatment research. It will be necessary to develop robust representations of the totality of outcomes for TB treatment, which will require collaboration between methodologists and clinicians. While different research questions may require incorporation of different outcomes measures into a prioritized outcome list, maintaining a level of standardization across similar studies is desirable to allow comparisons across studies and to inform further research. In addition, work is needed to determine the robustness of these methods when applied to the outcome distributions observed for TB treatment, including assessing the sensitivity of the results to changes in the outcome prioritization scheme.

If successfully applied, prioritized outcomes approaches for TB research could prove informative to clinicians by enabling an objective method for weighing risk and benefits against each other. Currently, clinicians presented with separate comparisons of safety and toxicity are required to make subjective risk–benefit assessments, and one unintended result of this situation may be overly conservative use of drugs (especially new drugs) because of a subjective emphasis on safety risks over treatment benefits. For example, using regimens with more drugs may reduce the risk of mortality and relapse [20,21], but the additional side effects attributed to the use of an additional drug may be deemed to outweigh any improvement of treatment outcomes. In addition, the new drugs bedaquiline and delamanid have recently been approved for treatment of MDR-TB, but uptake has been conservative in part because of concerns about the possibility of severe adverse events including death, even though these events were rare in clinical trials [22,23]. As illustrated in our example, assessing a totality of outcomes could be particularly useful in informing these types of decisions.

As in any methodological approach to comparing treatments, this strategy is not without its challenges and limitations. One major consideration is how to create the ordered categories. There are several ways to elicit the prioritized outcomes, and some suggestions are provided in Section 3 of this paper. There is also the possibility that a significantly inferior treatment with respect to the primary clinical outcome will come out better due to improvements in less important outcomes. A possibly effective way to address this is to use a “partial credit” strategy that would exclude a potentially influential but less important outcome in the prioritization scheme, and then directly assign its influence depending on the resulting distribution of the ordinal outcome [24]. Another way is to perform sensitivity analyses on the individual components, as is done with other composite outcomes. On the whole, despite the above mentioned limitations, one should not be deterred from using an approach based on totality of outcomes since it can potentially offer added value to the conventional analyses used in TB research.

As a research community, we should start by using this approach on available datasets in order to understand its behavior in the context of TB studies. The results of this preliminary work will not only help to develop these methods for TB outcomes

research but also equip us with knowledge of the assumptions we must make if we are to design future TB clinical trials using this new paradigm.

Funding

This work received funding support from the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under award number [UM1 AI068616](#) (IMPAACT SDMC).

Authors contribution

Montepiedra G. and Yuen C. conceived the manuscript idea and wrote the first draft of the manuscript. Evans S. and Rich M. contributed to interpretation and manuscript writing. All authors read, reviewed and approved the final version of the paper.

Acknowledgments

We are grateful to Dean Follmann for his comments and suggestions during the writing of this manuscript.

References

- [1] Definitions and reporting framework for tuberculosis 2013 revision, Geneva, Switzerland: World Health Organization; 2013. [WHO/HTM/TB/2013.2].
- [2] Guidelines for the programmatic management of drug-resistant tuberculosis, Geneva, Switzerland: World Health Organization; 2008. [WHO/HTM/TB/2008.402].
- [3] Bastard M, Bonnet M, du Cros P, Khamraev AK, Hayrapetyan A, Kimenye K, et al. Revised definitions of multidrug-resistant tuberculosis treatment outcomes: closer to the reality? *Am J Respir Crit Care Med* 2015;191(3):355–8.
- [4] Chuang-Stein C, Mohberg NR, Sinkula MS. Three measures for simultaneously evaluating benefits and risks using categorical data from clinical trials. *Stat Med* 1991;10(9):1349–59.
- [5] Follmann D, Wittes J, Cutler JA. The use of subjective rankings in clinical trials with an application to cardiovascular disease. *Stat Med* 1992;11(4):427–37.
- [6] Hallstrom P, Litwin PE, Weaver WD. A method of assigning scores to the components of a composite outcome: an example from the MITI trial. *Controlled Clin Trials* 1992;13(2):148–55.
- [7] Neaton JD, Wentworth DN, Rhame F, Hogan C, Abrams DI, Deyton L. Considerations in choice of a clinical endpoint for AIDS clinical trials. Terry Beinr community programs for clinical research on AIDS (CPCRA). *Stat Med* 1994;13(19–20):2107–25.
- [8] Bjorling LE, Hodges JS. Rule-based ranking schemes for antiretroviral trials. *Stat Med* 1997;16(10):1175–91.
- [9] Follmann DA. Regression analysis based on pairwise ordering of patients' clinical histories. *Stat Med* 2002;21(22):3353–67.
- [10] Buyse M. Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Stat Med* 2010;29(30):3245–57.
- [11] Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J* 2012;33(2):176–82.
- [12] Claggett B, Wei LJ, Pfeffer MA. Moving beyond our comfort zone. *Eur Heart J* 2013;34(12):869–71.
- [13] Rauch G, Jahn-Eimermacher A, Brannath W, Kieser M. Opportunities and challenges of combined effect measures based on prioritized outcomes. *Stat Med* 2014;33(7):1104–20.
- [14] Claggett B, Tian L, Castagno D, Wei LJ. Treatment selections using risk–benefit profiles based on data from comparative randomized clinical trials with multiple endpoints. *Biostatistics* 2015;16(1):60–72.
- [15] Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH, et al. Desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR). *Clin Infect Dis* 2015;61(5):800–6.
- [16] Dalkey NC. The Delphi method: an experiment study of group opinion. Monica Santa, editor. CA: Rand; 1969.
- [17] Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. *Pract Assess Res Eval* 2007;17(10):1–8.
- [18] Evans SR, Follmann D. Fundamentals and innovation in antibiotic trials. *Stat Biopharm Res* 2015;7(4):331–6.
- [19] Agresti A. Analysis of ordinal categorical data. Hoboken NJ, editor. 2nd ed. Wiley & Sons, Inc; 2010.
- [20] Franke MF, Appleton SC, Mitnick CD, Furin JJ, Bayona J, Chalco K, et al. Aggressive regimens for multidrug-resistant tuberculosis reduce recurrence. *Clin Infect Dis* 2013;56(6):770–6.
- [21] Mitnick CD, Franke MF, Rich ML, Alcantara Viru FA, Appleton SC, Atwood SS, et al. Aggressive regimens for multidrug-resistant tuberculosis decrease all-cause mortality. *PLoS One* 2013;8(3):e58664.
- [22] Diacon AH, Pym A, Grobusch MP, de los Rios JM, Gotuzzo E, Vasilyeva I, et al. Multidrug-resistant tuberculosis and culture conversion with bedaquiline. *N Engl J Med* 2014;371(8):723–32.
- [23] Gler MT, Skripconoka V, Sanchez-Garavito E, Xiao H, Cabrera-Rivero JL, Vargas-Vasquez DE, et al. Delamanid for multidrug-resistant pulmonary tuberculosis. *N Engl J Med* 2012;366(23):2151–60.
- [24] Evans SR. Benefit-risk and pragmatism: use outcomes to analyze patients rather than patients to analyze outcomes. In: Proceedings of Harvard catalyst seminar, Boston, MA; 2016.