Robert A. Weinstein, Section Editor

Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR)

Scott R. Evans,[1] Daniel Rubin,[2] Dean Follmann,[3] Gene Pennello,[4] W. Charles Huskins,[5] John H. Powers,[6,7] David Schoenfeld,[8] Christy Chuang-Stein,[9] Sara E. Cosgrove,[10] Vance G. Fowler Jr,[11] Ebbing Lautenbach,[12] and Henry F. Chambers[13]

[1]Harvard University, Boston, Massachusetts; [2]Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, [3]National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, and [4]Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, Maryland; [5]Division of Pediatric Infectious Diseases, Mayo Clinic, Rochester, Minnesota; [6]George Washington University School of Medicine, Washington D.C.; [7]Leidos Biomedical Research in support of the Division of Clinical Research, National Institutes of Health, Bethesda, Maryland; [8]Massachusetts General Hospital, Boston; [9]Pfizer, Inc, Groton, Connecticut; [10]Johns Hopkins University, Baltimore, Maryland; [11]Duke University, Durham, North Carolina; [12]University of Pennsylvania, Philadelphia; and [13]University of California, San Francisco

(See the Editorial Commentary by Molina and Cisneros on pages 807–8.)

Clinical trials that compare strategies to optimize antibiotic use are of critical importance but are limited by competing risks that distort outcome interpretation, complexities of noninferiority trials, large sample sizes, and inadequate evaluation of benefits and harms at the patient level. The Antibacterial Resistance Leadership Group strives to overcome these challenges through innovative trial design. Response adjusted for duration of antibiotic risk (RADAR) is a novel methodology utilizing a superiority design and a 2-step process: (1) categorizing patients into an overall clinical outcome (based on benefits and harms), and (2) ranking patients with respect to a desirability of outcome ranking (DOOR). DOORs are constructed by assigning higher ranks to patients with (1) better overall clinical outcomes and (2) shorter durations of antibiotic use for similar overall clinical outcomes. DOOR distributions are compared between antibiotic use strategies. The probability that a randomly selected patient will have a better DOOR if assigned to the new strategy is estimated. DOOR/RADAR represents a new paradigm in assessing the risks and benefits of new strategies to optimize antibiotic use.

Keywords. DOOR; RADAR; antibiotic use strategies.

Overuse of antibiotics is common, with estimates of 25%–75% of antibiotic use being unwarranted in acute care hospitals [1] and long-term-care facilities [2]. It is the major driver of the emergence of antibiotic resistance [3, 4] and is associated with adverse events such as *Clostridium difficile* infection. Antibiotic stewardship programs seek to optimize the use of antibiotics to limit these consequences; however, evidence regarding the efficacy and safety of various antibiotic use strategies (eg, shorter courses of therapy, use of narrow-spectrum agents) remains limited. Designing clinical trials to evaluate the effectiveness of such approaches is challenging for several reasons, including misleading outcomes, complexities of noninferiority trials, and the lack integration of benefits and harms. Herein we discuss issues with common approaches and introduce alternative methodology.

COMMONLY USED OUTCOMES CAN BE MISLEADING

Studies typically measure hospital days, intensive care unit days, and antibiotic use (a surrogate for future antimicrobial resistance). Fewer days are interpreted as better outcomes, but this can be misleading due to competing

Received 1 February 2015; accepted 13 May 2015; electronically published 25 June 2015.

Correspondence: Scott R. Evans, PhD, MS, FXB-513, CBAR/Biostatistics, Harvard School of Public Health, 651 Huntington Ave, Boston, MA 02115 (evans@sdac.harvard.edu).

Clinical Infectious Diseases® 2015;61(5):800–6

© The Author 2015. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

DOI: 10.1093/cid/civ495

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

risks (eg, a hospital stay may be brief due to mortality). Many (particularly nonrandomized) studies have utilized days of antibiotic therapy with, for example, 1000 hospital-days as an outcome, with a smaller value interpreted as a more desirable result. But this measure could be made smaller by increasing hospital stay.

COMPLEXITIES OF NONINFERIORITY TRIALS: QUESTIONING THE QUESTION

The commonly used noninferiority trial design with respect to clinical outcome (eg, cure) does not address the relevant question regarding whether one approach is better than another. The complexities associated with noninferiority trials are documented [5–9]. Compared with superiority trials, they are more prone to biases and manipulation, resulting in lower scientific integrity. Their validity relies upon several foundational assumptions.

The noninferiority margin is generally selected to ensure that a noninferiority result would (1) imply preservation of some of the effect that the control has historically displayed (ie, vs placebo), and (2) rule out with reasonable confidence clinically important levels of inferiority (using confidence intervals) so that clinical application is ethical and clinically acceptable. In a setting of evolving resistance, the constancy assumption (ie, the effectiveness of the control is unchanged) may not hold, creating challenges in selecting a margin that ensures preservation of the control effect. Even if such a margin is identified, it may be unconvincing to the medical community and not evidence-based from a clinical importance perspective.

Li et al discussed 7 anti-infective trials that aimed to evaluate noninferiority of placebo to antibiotic therapy for treatment of sinusitis, otitis media, and acute bacterial exacerbation of chronic bronchitis in an attempt to evaluate whether antibiotic use was unnecessary for these diagnoses [10]. Here the noninferiority margin clearly cannot be derived based on the concept of preservation of the effect, as the goal of the trial is to demonstrate that there is no effect to retain (ie, placebo cannot be superior to placebo).

In noninferiority trials, there is an incentive to reduce assay sensitivity such that strategies appear similar by diluting effects through subtle choices about design and conduct, including entry criteria (eg, the presence of specific pathogens), endpoint selection and timing, adherence, prior/concomitant therapy, or loss to follow-up. Blinding provides muted protection from bias, as a blinded investigator can skew results toward similarity by assigning similar response ratings for all participants [5, 11].

Noninferiority trials raise ethical dilemmas [12]. The null hypothesis is that the experimental strategy is inferior, viewed by some as a violation of the equipoise necessary for randomization. Patients may not be informed of this, contradicting the principles of informed consent and justice in clinical research.

Noninferiority trials can be unattractive for patients and clinicians. Why would patients volunteer to risk randomization to a strategy that *may be* as good as an existing, proven medical alternative but is not hypothesized to be better? Why not simply opt for the proven alternative?

Last, noninferiority trials frequently have large and impractical sample sizes that jeopardize feasibility and strain resources.

EVALUATING BENEFITS AND HARMS AND PATIENT-LEVEL INTERPRETATION

The synthesis of clinical and antibiotic use outcomes requires careful thought. Trial results show that some patients may benefit while some may experience harm (adverse effects). Evaluating the association between these outcomes is important for understanding the overall effects for individual patients. If the group of patients experiencing harm and the group of patients experiencing benefit are largely disjointed, then it is important to identify ways to distinguish between these 2 groups prior to applying the intervention. The value of the intervention will depend on the ability to identify and avoid intervention use in those harmed, while targeting the intervention toward those who benefit. However, if the 2 groups are largely overlapping, then an assessment is needed to determine the net effect (ie, whether the benefits outweigh the harms). The traditional approach to the analysis of trials is to separately analyze each endpoint (eg, treatment success, significant adverse event). However, this practice cannot distinguish between the 2 scenarios described above and thus does not optimally evaluate the distribution of the totality of the effects on individual patients, a critical element for medical decision making. Aggregation of the total disease burden and experience of each patient is needed.

DOOR AND RADAR

Given the inability of standard approaches to address these challenges, new methodologies are needed. The desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR) methods have been developed for this purpose.

DOOR

As the name suggests, DOOR is a ranking of all trial participants with respect to the desirability of their overall outcome. During the analyses of a clinical trial, the distributions of DOORs are compared between strategies. The construction of DOOR begins with defining an (ordinal) overall clinical outcome.

Overall Clinical Outcome

The overall clinical outcome is based on a *longitudinal snapshot* of the experience of the individual patient during the course of the trial, analogous to a *discharge review* or *exit examination* frequently conducted during hospital discharge, but now applied to the clinical trial setting. Each patient experience is

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

characterized according to the overall clinical outcome carefully constructed on the basis of important clinical outcomes (ie, benefits, harms, and possibly quality of life). This provides a comprehensive synthesis of the results for an individual patient. Outcomes are used to analyze patients rather than using patients to analyze outcomes.

Consider the following generic example, where the overall clinical outcome has 5 mutually exclusive hierarchical levels in descending order of desirability:

- Clinical benefit (patient symptoms/function) without adverse effects (AEs)
- Clinical benefit with some AEs
- Survival without clinical benefit or AEs
- Survival without clinical benefit but with AEs
- Death

All trial participants are categorized according to the overall clinical outcome. Participants in different categories have clinically relevant differences on overall clinical outcome. Participants in the same category have similar overall clinical outcomes. The number and definition of categories is tailored to the clinical disease of interest (eg, more levels could be created based on AE types/severity and emergence of resistance; or the death category could be dichotomized into early vs late death). In the example above, the order of the second and third categories may be reversed or collapsed into a single category in some instances. For example, it may be desirable to construct an overall clinical outcome such that a participant who is treated and cured for a urinary tract infection (UTI) but develops a severe *C. difficile* infection requiring colectomy is classified into a worse overall outcome than a participant who is not originally cured from the UTI but improves with no AEs.

Consensus regarding the definition of the overall clinical outcome is critical. Construction can be challenging and requires critical thought (suggestions for construction are provided in Table 1). The resulting definition should be defined clearly in the trial protocol. As the construction may involve subjective components (eg, clinical cure), the use of double-blind designs or blinded adjudication committees should be considered.

RADAR

RADAR is a version of DOOR tailored for trials that compare strategies to optimize antibiotic use. RADAR synthesizes clinical patient outcomes with antibiotic use outcomes under the principle that less antibiotic use is better but cannot be at the expense of clinical outcomes. RADAR utilizes a superiority trial design, evaluating whether a new strategy is better than a current strategy when considering all of the important outcomes (benefits and harms, duration of antibiotic use) when logically prioritized. New strategies have value if they are superior to current strategies when all effects are considered together.

In RADAR, all trial participants are assigned a DOOR. The DOOR is constructed using a 2-step process: (1) categorization of all patients into an overall clinical outcome, and (2) ranking participants in the trial using 2 rules:

1. When ranking the outcomes of 2 patients with different overall clinical outcomes, the patient with a better overall clinical outcome receives a higher rank.
2. When ranking the outcomes of 2 patients with the same overall clinical outcome, the patient with a shorter duration of antibiotic use receives a higher rank.

Thus clinical outcome trumps the duration of antibiotic use (ie, a patient with a worse clinical outcome cannot have a higher rank than a patient with a better clinical outcome, regardless of the duration of antibiotic use).

RADAR analyzes the *pragmatic strategy* of interventions as applied in practice. Strategy evaluation is the most clinically

Table 1. Suggestions for Overall Clinical Outcome Construction

- Consider:
  (a) The general overall (benefit and risk) patient-level clinical outcomes that have differing levels of importance, and
  (b) How patients tend to cluster themselves in terms of overall clinical outcomes (ie, categories may be naturally apparent).
- Identify and prioritize important clinical outcomes including efficacy, safety, and quality of life. Some factors can be viewed as equally important.
- Use "all-cause" outcomes when formulating the response in randomized studies. Patients, not specific outcomes, are being evaluated. Causality is evaluated by a contrast of the randomized strategies rather than judged "relatedness" to the intervention or disease. If an outcome is unrelated to treatment, then it will occur with similar frequency between randomized arms. If an outcome occurs differentially between arms, then it is related to treatment.
- Consider using outcomes that are a function of the patient, or standardize the criteria for clinical decisions (eg, duration of hospital stay), to eliminate or reduce variation induced by clinician decision. A fundamental tenet in clinical trials is to minimize variation as this provides the best opportunity to identify intervention effects if they exist. Although some measures (eg, change of therapy) can be objectively measured, they are partly a function of clinician decision in addition to patient outcome. This adds another source of variation (ie, due to the clinician). Patients may switch therapy because of clinical failure but should not necessarily be considered a clinical failure because they switch therapy. Clinician decisions can be used as a surrogate for patient responses when detailed patient responses are unavailable.
- Use endpoints that are clinically meaningful and that are measures of how patients feel, function, or survive. Limit the use of biomarkers unless they are thoroughly validated as surrogates for clinical patient response. Many currently utilized biomarkers have not been validated as adequate surrogates. If patient responses are observable within reasonable time frames, then surrogates may be unnecessary.

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

relevant evaluation, as the decision regarding strategy initiation must be made prior to knowledge of subsequent patient results that may necessitate adjustments to therapy. Adherence (ie, observed antibiotic use) is incorporated into the DOOR. Consider a trial designed to compare the outcomes of strategies of 5 vs 10 days of antibiotics. A participant who is randomized to the 5-day strategy, but receives 10 days of antibiotics, is evaluated for overall clinical response as part of the 5-day strategy, consistent with the intention-to-treat principle, even though the observed treatment is not as intended. But the DOOR of the participant considers that the participant received 10 days of therapy.

RADAR utilizes the duration of antibiotic use in days. However, other measures of antibiotic exposure (eg, the number of doses, intensity, antibiotic class, intravenous vs oral) could also be incorporated as measures of selected pressure for resistance. For example, a ranking scheme that is constructed to prioritize broad-spectrum use (less use implies better rank), and then narrow-spectrum antibiotic use, could break ties within similar broad-spectrum use.

Illustration

DOOR/RADAR is illustrated with the following example. Consider a randomized trial designed to compare a new strategy (13 trial participants, denoted A–M) vs a control strategy (13 participants, denoted N–Z). The experience of each of the 26 trial participants is categorized into 3 mutually exclusive categories: success without AE, success with AE, and failure according to an overall clinical outcome. The duration of antibiotic use is also observed for each participant.

Data for each of these 26 participants and DOORs are shown in Table 2. The new and control strategies have similar distributions of the overall clinical outcome. However, the new strategy generally results in fewer days of antibiotic use; thus, the new strategy has higher DOORs. The probability of a better DOOR

Table 2. Response Adjusted for Duration of Antibiotic Risk Illustration: Participant Data and Data Summaries

Participant	Treatment Arm	Overall Clinical Outcome[a]	Days of Antibiotic Use	DOOR	No. of Control Participants (n = 13) With a Lower DOOR
A	New	2	5	11	9
B	New	1	3	1	13
C	New	1	4	2	13
D	New	2	4	10	9
E	New	3	3	19	4
F	New	2	3	9	9
G	New	3	5	21	4
H	New	3	4	20	4
I	New	1	7	5	12
J	New	3	8	23	3
K	New	2	6	12	9
L	New	1	5	3	13
M	New	2	8	14.5	7.5
N	Control	3	12	26	Sum = 109.5
O	Control	2	7	13	
P	Control	1	9	7	
Q	Control	2	8	14.5	
R	Control	3	6	22	
S	Control	2	11	18	
T	Control	1	10	8	
U	Control	2	9	16	
V	Control	3	9	24	
W	Control	1	6	4	
X	Control	1	8	6	
Y	Control	2	10	17	
Z	Control	3	10	25	

The probability of a better DOOR for a randomly selected participant from the new strategy compared with the old strategy is the number of between-treatment pairwise comparisons in which the new treatment has a higher DOOR than the control (109.5), divided by the total number of possible pairwise comparisons (169), resulting in 64.8% (95% confidence interval, 57%–71%).

Abbreviations: AE, adverse effects; DOOR, desirability of outcome ranking.

[a] Overall clinical outcome coding: 1, success without AE; 2, success with AE; 3, failure.

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

for a randomly selected participant from the new strategy vs the old strategy is 64.8% (95% confidence interval, 57%–71%).

ANALYSES AND SAMPLE SIZE

During analyses, the distributions of DOORs are compared between strategies. The probability that a randomly selected patient will have a better DOOR if assigned to the new strategy vs the control strategy is estimated using a confidence interval. If there is no difference in DOOR distributions between the 2 strategies, then the probability will be near 50%. If there is no between-strategy difference in clinical outcomes but there is a reduction in antibiotic use with the new strategy, then the probability will be >50%. If the new strategy offers an advantage in clinical outcomes and reduction in antibiotic use, then the probability will be even higher.

The sample-size calculation is based on a superiority test with the following hypotheses:

- Null: No difference in DOOR.
- Alternative: The new strategy has a higher DOOR (ie, the probability that a randomly selected patient will have a better DOOR if assigned to the new strategy vs the control strategy is >50%).

Similar to alternative hypotheses for other tests, the selected magnitude of superiority to detect is based on the minimum clinically important difference concept. DOOR/RADAR may have an intuitive relative attractiveness for selecting the minimum important difference due to the patient-level nature of the hypothesis. Trials can be sized via standard software (eg, EAST) using rank-based methods (eg, the Wilcoxon–Mann–Whitney test) [13]. Sample size can be further evaluated using simulation, using assumptions about the overall clinical outcome rates and the distribution of antibiotic use within outcome categories.

A concern with RADAR is whether a decrement in the overall clinical outcome or a component thereof (an important outcome) could be offset by a large improvement in the reduction in antibiotic use (an outcome of comparatively lesser importance), despite DOOR directly incorporating the relative priorities of these endpoints into the ranking strategy. As with other composite endpoints, the advantage of a strategy on the DOOR analysis does not necessarily imply an advantage on all of the components; in fact, disadvantages on specific components are possible. Thus, examination of the effects on the overall clinical outcome and each component is standardly conducted via sensitivity analyses. Analyses that dampen the impact of antibiotic use are also possible. If a research team wishes to size the trial to evaluate a specific component outcome, then the sample size may need to be increased, as the sample size required for the DOOR analyses may be smaller than that required for evaluation of individual component outcomes.

EXAMPLE

We illustrate RADAR in the design of a trial being developed by the Antibacterial Resistance Leadership Group [14] and funded by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health.

The Short-Course Outpatient Therapy for Community-Acquired Pneumonia in Children trial is a randomized, double-blind, placebo-controlled clinical trial to compare the strategy of a short (5-day) vs standard (10-day) course of β-lactam antibiotic therapy in children aged 6 months to <6 years with community-acquired pneumonia.

The trial was initially proposed as a noninferiority trial with a 5% noninferiority margin and a primary endpoint of treatment failure at the test-of-cure visit 14 days after initiation of therapy. Based on the assumption of a 5% treatment failure rate in the standard arm, using a 1-sided $\alpha = .025$ test, a sample size of 800 trial participants (400 per arm) was proposed and would have provided 90% power to detect noninferiority. Reviewers of the proposal had primary concerns regarding the appropriate noninferiority margin and the feasibility of enrolling the required sample size.

To address these concerns, the trial was redesigned as a superiority trial using RADAR. An overall clinical outcome of 8 mutually exclusive hierarchical levels was developed (Table 3). Each participant will be classified into these levels and then assigned a DOOR.

The sample size was based on a superiority test with an alternative hypothesis that there is a 60% probability that a patient assigned to the short-course strategy will have a higher DOOR than if the patient were assigned to the standard-course strategy.

Table 3. Overall Clinical Outcome for the SCOUT-CAP[a] Trial (From Most to Least Desirable)

1.	Survival; adequate clinical response; no adverse events
2.	Survival; adequate clinical response; mild adverse event(s)
3.	Survival; adequate clinical response; moderate adverse event(s)
4.	Survival; adequate clinical response; severe adverse event(s)
5.	Survival; inadequate clinical response without additional emergency department or clinic visit or hospitalization
6.	Survival; inadequate clinical response with additional emergency department or clinic visit but without hospitalization; any grade of adverse event
7.	Survival; inadequate clinical response with hospitalization; any grade of adverse event
8.	Death

Adequate clinical response was defined based on absence of all of the following as assessed on day 11–14 after initiation of therapy: (1) fever unless related to a new process that is unrelated to the prior diagnosis of pneumonia, (2) tachypnea, (3) increased work of breathing (retractions, nasal flaring, grunting), and (4) a medically attended visit to an emergency department/clinic or hospitalization for persistent or worsening pneumonia at any time after randomization. This is a draft version from a developing clinical trial.

[a] Short-Course Outpatient Therapy for Community-Acquired Pneumonia in Children.

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

In total, 360 participants (180 per strategy) provide 90% power to detect superiority using a 2-sided $\alpha = .05$ (EAST) [13]. The required sample size represents more than a 50% decrease compared with the original noninferiority design. Of note, a superiority trial designed to detect a difference between 90% and 95% would require 870 total participants to have 80% power with a 2-sided $\alpha = .05$.

CONCLUSIONS

DOOR is a methodology designed to address several challenges in clinical trials. RADAR is a version of DOOR that can be applied to pragmatic stewardship trials comparing strategies to optimize antibiotic use. Together, DOOR/RADAR (1) evaluate the clinically relevant question of superiority of the new strategies based on consideration of all consequences, avoiding the complexities associated with a noninferiority design; (2) incorporate competing risks (eg, death) and adherence as part of the outcome within the context of other clinical data, making results more interpretable; (3) directly evaluate the association between benefits and harms of antibiotic use; (4) allow for patient-level interpretation by using outcomes to analyze patients rather than using patients to analyze outcomes; and (5) potentially reduce sample sizes, making trials less costly and more feasible.

DOOR/RADAR can also be used to supplement noninferiority analyses either as co-primary analyses, as multiple primary analyses, or as secondary analyses with an appropriate error-control strategy for multiplicity. If a noninferiority trial is conducted with a 10% noninferiority margin and the relevant confidence interval bound for the treatment difference is 8%, the trial is able to rule out inferiority of 10% with reasonable confidence. Although the criteria for noninferiority were met, inferiority of up to 8% cannot be ruled out with reasonable confidence (ie, up to 8% inferiority is consistent with the data). This result may be unconvincing to clinicians and patients who are unwilling to take such a risk associated with a new strategy. Whereas noninferiority trials provide an evaluation based on an important endpoint, DOOR/RADAR provides a comprehensive evaluation of superiority based on all important endpoints. Demonstration of superiority with respect to DOOR may be illuminating and provide assurance and incentive in the decision to opt for a new strategy.

DOOR/RADAR has limitations. Numerous ordinal levels, adding to study complexity, may be required for complicated outcomes. Developing a predetermined, consensus ordinal ranking strategy could be difficult. Ordering/ranking can become subjective, as evidence-based criteria may be lacking (eg, criteria used to define broad- vs narrow-spectrum antimicrobials may not be validated). Application to typical quasi-experimental stewardship studies may be limited by the sheer number of variables needed to describe different agents, durations, and AEs. AEs and outcomes related to emergence of resistance will depend on the laboratory tests that the patient happens to have completed. These issues notwithstanding, DOOR/RADAR represents a novel approach to study design in assessing the risks and benefits of new strategies to optimize antibiotic use.

Notes

Disclaimer. This article should not be construed to represent the US Food and Drug Administration's views or policies. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH).

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases of the NIH (award number UM1AI104681) and in part with federal funds from the National Cancer Institute, NIH (contract number HHSN261200800001E).

Potential conflicts of interest. W. C. H. receives salary support for his work on the Short-Course Outpatient Therapy for Community-Acquired Pneumonia in Children trial under the Antibacterial Resistance Leadership Group (award number UM1AI104681). V. G. F. has served on the board of Merck; has served as a consultant for Pfizer, Novartis, Galderma, Novadigm, Durata, Debiopharm, Genentech, Achaogen, Affinium, The Medicines Co, Cerexa, Tetraophase, Trius, MedImmune, Bayer, Theravance, Cubist, and Genentech; has been an employee of Duke University Medical Center; has received grants from NIH, MedImmune, Forest/Cerexa, Pfizer, Merck, Advanced Liquid Logics, Theravance, Novartis, and Cubist; holds patents from National Center for Genome Resources; has received royalties from UpToDate; and has received payment for development of educational presentations from NIH, MedImmune, Forest/Cerexa, Pfizer, Merck, Advanced Liquid Logics, Theravance, Novartis, and Cubist. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Cosgrove SE, Seo SK, Bolon MK, et al. Evaluation of post-prescription review and feedback as a method of promoting rational antimicrobial use: a multicenter intervention. Infect Control Hosp Epidemiol **2012**; 33:374–80.
2. Peron EP, Hirsch AA, Jury LA, Jump RL, Donskey CJ. Another setting for stewardship: high rate of unnecessary antimicrobial use in a Veterans Affairs long-term care facility. J Am Geriatr Soc **2013**; 61:289–90.
3. Paterson DL. "Collateral damage" from cephalosporin or quinolone antibiotic therapy. Clin Infect Dis **2004**; 38(suppl 4):S341–5.
4. Rice LB. The clinical consequences of antimicrobial resistance. Curr Opin Microbiol **2009**; 12:476–81.
5. Snappin SM. Noninferiority trials. Curr Control Trials Cardiovasc Med **2000**; 1:19–21.
6. Powers JH, Cooper CK, Lin D, Ross DB. Sample size and the ethics of non-inferiority trials. Lancet **2005**; 366:24–5.
7. Fleming TR. Current issues in non-inferiority trials. Stat Med **2008**; 27:317–32.
8. Powers JH. Noninferiority and equivalence trials: deciphering "similarity" of medical interventions. Stat Med **2008**; 27:343–52.
9. Evans S. Noninferiority clinical trials. Chance **2009**; 22:53–8.
10. Li Y, He Y, Sheng Y, et al. Systematic evaluation of non-inferiority and equivalence randomized trials of anti-infective drugs. Expert Rev Anti Infect Ther **2013**; 11:1377–89.

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021

11. Oczkowski SJ. A clinician's guide to the assessment and interpretation of noninferiority trials for novel therapies. Open Med **2014**; 8:e67–372.
12. Djulbegovic B, Clarke M. Scientific and ethical issues in equivalence trials. JAMA **2001**; 285:1206–8.
13. Noether GE. Sample-size determination for some common nonparametric tests. J Am Stat Assoc **1987**; 82:645–7.
14. Chambers HF, Bartlett JG, Bonomo RA, et al. Antibacterial resistance leadership group: open for business. Clin Infect Dis **2014**; 58:1571–6.

Downloaded from https://academic.oup.com/cid/article/61/5/800/304719 by Yale University Law School user on 17 February 2021