Clinical Infectious Diseases

MAJOR ARTICLE

Good Studies Evaluate the Disease While Great Studies Evaluate the Patient: Development and Application of a Desirability of Outcome Ranking Endpoint for *Staphylococcus aureus* Bloodstream Infection

Sarah B. Doernberg,[1] Thuy Tien Tram Tran,[2] Steven Y. C. Tong,[3,4] Mical Paul,[5,6] Dafna Yahav,[7,8] Joshua S. Davis,[4,9] Leonard Leibovici,[8,10] Helen W. Boucher,[11] G. Ralph Corey,[12] Sara E. Cosgrove,[13] Henry F. Chambers,[1] Vance G. Fowler,[12] Scott R. Evans,[2] and Thomas L. Holland[12]; for the Antibacterial Resistance Leadership Group

[1]Department of Internal Medicine, Division of Infectious Diseases, University of California, San Francisco; [2]Biostatistics Center, Department of Epidemiology and Biostatistics, Milken Institute School of Public Health, George Washington University, D.C.; [3]Victorian Infectious Disease Service, the Royal Melbourne Hospital, and the University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Victoria, Australia; [4]Global and Tropical Health Division, Menzies School of Health Research, Charles Darwin University, Northern Territory, Australia; [5]Institute of Infectious Diseases, Rambam Health Care Campus, and [6]Faculty of Medicine, Technion, Israel Institute of Technology, Haifa, [7]Unit of Infectious Diseases, Rabin Medical Center, Beilinson Hospital, Petah Tikva, and [8]Sackler Faculty of Medicine, Tel Aviv University, Ramat-Aviv, Israel; [9]Department of Infectious Diseases, John Hunter Hospital, Newcastle, New South Wales, Australia; [10]Department of Medicine E, Rabin Medical Center, Beilinson Hospital, Petah Tikva, Israel; [11]Tufts Medical Center, Infectious Diseases, Boston, Massachusetts; [12]Duke University Medical Center, Durham, North Carolina; and [13]Department of Antimicrobial Stewardship, Johns Hopkins University, Baltimore, Maryland

Background. Desirability of outcome ranking (DOOR) is an innovative approach in clinical trials to evaluate the global benefits and risks of an intervention. We developed and validated a DOOR endpoint for *Staphylococcus aureus* bloodstream infection (BSI) through a survey to infectious diseases clinicians and secondary analysis of trial data.

Methods. We administered a survey of 20 cases of *S. aureus* BSI, asking respondents to rank outcomes by global desirability. Correlations and percentage of pairwise agreement among rankings were estimated to inform development of a DOOR endpoint, which was applied to 2 prior *S. aureus* BSI trials. The probability that a patient randomly assigned to experimental treatment would have a better DOOR ranking than if assigned to control was estimated. Results were also analyzed using partial credit, which is analogous to scoring an academic test, assigning 100% to the most desirable outcome, 0% to the least, and "partial credit" to intermediate ranks.

Results. Forty-two recipients (97%) completed the survey. The DOOR endpoint fitting these rankings ($r = 0.89$; 95% confidence interval, 0.67 to 0.94) incorporated survival plus cumulative occurrence of adverse events, cure, infectious complications, and ongoing symptoms. Tailored versions of this endpoint were applied to 2 *S. aureus* BSI trials, and both demonstrated no benefit of the experimental treatment using DOOR and partial credit analysis.

Conclusions. Using *S. aureus* BSI as an exemplar, we developed a DOOR endpoint that can be used as a template for development of DOOR endpoints for other diseases. Future trials can incorporate DOOR to allow for global assessment of patient experience.

Keywords. clinical trial; *Staphylococcus aureus*; bloodstream infection.

To demonstrate efficacy and safety, new antibiotics are typically evaluated in active comparator noninferiority (NI) phase 3 trials that use binary endpoints, such as cure vs no cure. While NI trials are vital mechanisms for regulatory approval, they generally do not provide practical information about which treatment strategy represents the best overall option—balancing efficacy and safety—for an individual patient [1–3]. Even when designed to demonstrate superiority, trials designed to evaluate whether one treatment is more "successful" than another still neglect to assimilate information about both benefits and harms in a way that allows for global evaluation of the 2 treatment options. Specifically, binary efficacy outcomes (eg, cure vs failure) do not evaluate the association between component outcomes (ie, efficacy, toxicity, quality of life); do not evaluate the cumulative nature of component outcomes; do not systematically incorporate the relative importance of the component outcomes; and do not address competing risks. As a result, these trials fail to mimic real-world treatment decision-making. When evaluating treatments, clinicians integrate and weigh efficacy, safety, and quality-of-life considerations in a nonbinary fashion, taking into account the treatments as well as patient preference. Desirability of outcome ranking (DOOR) is a novel approach to evaluation of the global benefits and risks of an intervention and provides more pragmatic information for medical decision-making that complements traditional trial evaluations [2]. DOOR effectively uses outcomes to analyze patient experience as a whole rather than as separate

Received 15 May 2018; editorial decision 27 August 2018; accepted 21 September 2018; published online October 12, 2018.

Correspondence: T. L. Holland, Duke University Medical Center, Box 102359, Durham, NC 27710 (thomas.holland@duke.edu).

Clinical Infectious Diseases® 2018;XX(XX):1–8

DOI: 10.1093/cid/ciy766

components. Because the DOOR approach simultaneously considers effectiveness and toxicity when establishing outcome, researchers can define the global superiority of an intervention in a single outcome. In addition to DOOR, a complementary strategy called "partial credit" permits preferential assignment of relative importance to the DOOR levels for a more personalized analysis [4].

Staphylococcus aureus bloodstream infection (BSI) is a serious, common infection without a defined optimal treatment strategy [5]. To date, only 2 high-quality randomized, controlled trials for *S. aureus* BSI have been performed [5–8]. Approximately 40% of patients with *S. aureus* BSI develop metastatic infectious complications, and approximately 10% of infections relapse. Even in the best case, treatment for *S. aureus* BSI involves extended durations of potentially toxic antibiotics, long-term intravenous access, and blood draws for drug monitoring [9]. A trial design of *S. aureus* BSI comparing management strategies using a DOOR endpoint would provide pragmatic and patient-centered information that clinicians need in order to make informed decisions about patient care.

Our primary objective in this study was to develop and evaluate a DOOR endpoint for *S. aureus* BSI to be used in future clinical trials. We first conducted a survey of infectious diseases (ID) clinician–scientists and clinical trialists using case vignettes to develop a novel DOOR endpoint. Next, we applied this novel endpoint and partial credit approach as a secondary analysis of 2 previously completed clinical trials [10, 11].

METHODS

Survey Instrument

Twenty adult *S. aureus* BSI patient profiles were constructed by a team of ID physicians (S. B.D., T. L. H., H. F. C.) to represent the range of experiences and outcomes observed in prior trials. The profiles described efficacy, adverse events (AEs), symptoms, and treatment adjustments during a theoretical trial comparing 2 treatment strategies. These profiles required respondents to weigh trade-offs regarding patient outcomes. A computerized survey platform (Qualtrics, Provo, UT) presented the profiles to respondents in random order. A total of 43 physician members of the National Institutes of Health Antibacterial Resistance Leadership Group (ARLG) were sent the survey and asked to rank profiles from best to worst based on global patient outcome (see Supplementary Methods). Survey outcomes are described in Table 1.

Survey Statistical Analyses

Descriptive statistics were calculated for the clinician rankings of each patient profile. The Wilcoxon signed rank test was used to conduct pairwise comparisons of patient rankings. Spearman correlation coefficients and the percentage of pairwise agreement among clinician rankings were estimated. A DOOR endpoint based on respondent consensus was developed using classification and regression tree and discussion with experts.

Secondary Analyses

Methods for the 2 clinical trials for which we applied the DOOR endpoint have been previously published [10, 11]. Briefly, CAMERA (Combination Antibiotics for Methicillin-Resistant *Staphylococcus aureus*)-1 was an open-label, multicenter, randomized trial of 60 adults with methicillin-resistant *S. aureus* (MRSA) BSI randomized to treatment with vancomycin plus flucloxacillin or vancomycin alone to test whether the addition of flucloxacillin would shorten bacteremia duration [10]. Secondary outcomes included mortality, prolonged bacteremia, recurrent bacteremia, intensive care unit admission or septic shock after randomization, grade 2 nephrotoxicity or hepatotoxicity, or new metastatic complications. The second trial was an open-label, multicenter, randomized, controlled NI trial of trimethoprim-sulfamethoxazole vs vancomycin for treatment of inpatients with severe MRSA infection [11]. For this study, we analyzed the subset of 91 patients with BSI. Because the available outcomes differed for each study, the DOOR developed from the survey data was tailored for each trial by group consensus (S. B. D., T. T. T. T., H. W. B., G. R. C., S. E. C., H. F. C., V. G. F., S. R. E., T. L. H.), with input from the original authors of each study (Table 1). We analyzed all randomized patients in accordance with the intention-to-treat principle. Trial participants were classified into the categories of the DOOR outcome; participants missing a DOOR component were assigned the worst ranking of all survivors with full data. The distribution of DOOR rankings was compared between treatment groups. We estimated the probability that a patient randomly assigned to the experimental treatment strategy would have a better DOOR ranking than if assigned to vancomycin (the control for both trials); a probability of 0.5, or 50%, indicates no difference between groups. The ordinal outcomes included in the DOOR endpoint were also analyzed using a partial credit strategy [4]. This approach is analogous to scoring an academic test, assigning 100% to the most desirable outcome, 0% to the least (eg, death), and "partial credit" to each intermediate DOOR rank. The contrast between treatment strategies can be compared by varying the partial credit assignment of the intermediate ranks, which allows personalized clinician/patient preferences to be incorporated.

RESULTS

DOOR Endpoint Development

Forty-two of 43 ID clinician–scientists (97%) responded to the survey. Respondents generally agreed on global ranking (median Spearman correlation r = 0.69; interquartile range, 0.60–0.77). Respondents tended to differentiate best and worst profile outcomes (ie, patients with none or several component outcome events), whereas those within the middle were more difficult to distinguish (Figure 1). Within this middle group, there were 3 major groupings of several patient profiles that appeared interchangeable. A de novo DOOR endpoint was generated (Table 2) to best fit respondent consensus rankings (r = 0.89; 95% confidence interval [CI], 0.67 to 0.94).

Table 1. Component Definitions for Secondary Analyses Based on Available Data

Component	Survey	Combination Antibiotics for Methicillin Resistant *Staphylococcus aureus*-1 Trial	Trimethoprim-Sulfamethoxazole Trial
Treatment failure	Lack of global resolution of all sites of *S. aureus* infection at the test of cure 8 weeks after randomization	At least 1 of the following: - Relapse (positive blood culture for MRSA isolated ≥48 hours after a negative blood culture) - Readmission attributable to MRSA within 90 days - Persistent bacteremia (blood cultures remaining positive ≥ day 5 after randomization)	Positive blood culture for MRSA on or after day 5 from randomization
Infectious complications	At least 1 of the following: - Development of drug resistance - Newly identified metastatic focus of infection - Persistent or resistant *S. aureus* BSI[a]	At least 1 of the following post-randomization: - Endocarditis - Osteoarticular infection - Infection-related stroke - Deep abscess - Septic shock - Intensive care unit admission	New post-randomization MRSA-related diagnosis
Ongoing symptoms	Ongoing moderate to severe pain or other localizing symptoms at the site of documented *S. aureus* infection that limit daily activities, with or without evidence of ongoing infection (eg, back pain at site of treated vertebral osteomyelitis)	Unavailable	Unavailable
Grade 4 AE [12, 13]	Any unfavorable or unintended sign, symptom, or disease temporally associated with the use of the antibiotic but with no judgment about causality; severity of AE was graded according to Department of Health and Human Services common terminology criteria for AEs	At least 1 of the following Grade 4 laboratory abnormalities: - Platelet count - Bilirubin - Albumin - Creatinine - alanine aminotransferase - gamma-glutamyl transferase - white blood cell count	At least 1 of the following: - New need for dialysis post-randomization - Discontinuation of medications due to life-threatening AE - Stevens-Johnson syndrome - Meeting "F" or "L" of the RIFLE (risk, injury, failure, loss, and end-stage kidney disease) criteria (eg, renal failure or renal loss)
Death	Until the end of the follow-up period (8 weeks after randomization)	90-day mortality	30-day mortality

Abbreviations: AE, adverse event; BSI, bloodstream infection; MRSA, methicillin-resistant *Staphylococcus aureus*.

[a]Persistent BSI was defined as blood cultures remaining positive ≥5 days after randomization, while relapsed BSI included patients who had negative blood cultures followed by positive blood cultures on or after day 5 [14]. A patient with persistent or relapsed BSI could still be cured if blood cultures were negative by test of cure.

Real-world Application I: Secondary Analysis of the CAMERA-1 Study

The tailored DOOR endpoint for this study included the following components (Table 1): mortality at 90 days, treatment failure, infectious complications, and grade 4 AEs during study follow-up up to 30 days. The 5 possible ranks were: (1) alive without treatment failure, infectious complications, or grade 4 AE; (2) alive with any 1 of treatment failure, infectious complications, or grade 4 AE; (3) alive with any 2 of treatment failure, infectious complications, or grade 4 AE; (4) alive with all 3 of treatment failure, infectious complications, or grade 4 AE; and (5) dead. There were no patients with rank 4 in the study (Figure 2). Patients who received vancomycin plus flucloxacillin had a 47% chance of having a better DOOR compared to those who received vancomycin alone (95% CI, 33% to 60%; Table 3). The confidence interval crosses 50%, where the chance of getting a better DOOR is essentially random; this result suggests that the global outcome for a patient getting combination therapy is no different than the global outcome for a patient receiving vancomycin alone.

Between-treatment differences were calculated for a range of partial credit scores between 0 and 100 assigned to rankings 2–4, with rank 1 (alive, no complications) assigned a score of 100 and rank 5 (death) assigned a score of 0. Partial credit scores were ranged such that 0 ≤ rank 4 ≤ rank 3 ≤ rank 2 ≤ 100. Table 3 displays a range of potential partial credit assignments with several exemplar scenarios. Scenario A represents the extreme example where survival is the only relevant outcome and is akin to a traditional mortality analysis. Death is assigned a score of 0 while all other ranks were assigned a score of 100. Scenario B represents the other extreme, where any unfavorable outcome (treatment failure, infectious complications, or grade 4 AE) is considered equal to death and assigned a score of 0. Scenarios C and D represent 2 intermediate partial credit scoring approaches. Supplementary Figure 1 demonstrates the distribution of all possible weight assignments, with specific scenarios shown in Supplementary Table 1. There were no statistically significant differences between treatment arms by any assigned partial credit strategy.

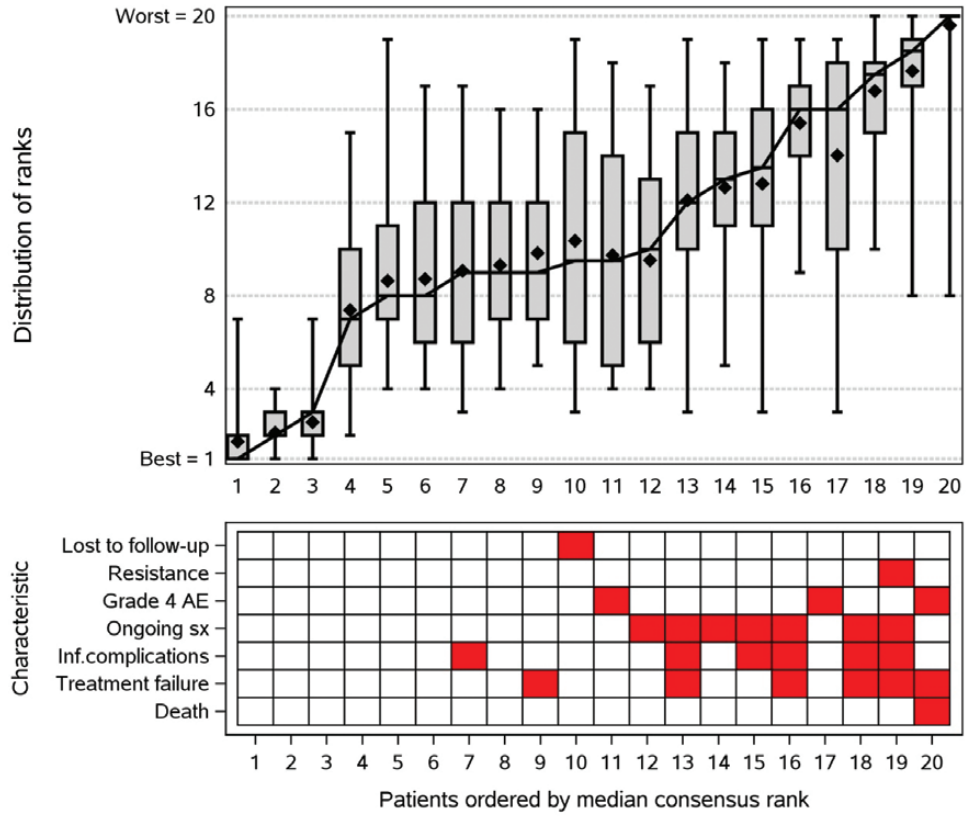

Figure 1. Boxplot of ranks by patient with patients ordered by median consensus clinician rank. Abbreviations: AEs, adverse events; sx, ongoing symptoms.

Real-world Application II: Secondary Analysis of Trimethoprim-Sulfamethoxazole vs Vancomycin for Treatment of MRSA

The DOOR for this study included the following components determined at day 30 after randomization (Table 1): death, treatment failure, infectious complications, and grade 4 AEs. The 5 possible ranks were similar to the CAMERA-1 rankings: (1) alive without treatment failure, infectious complications, or grade 4 AE; (2) alive with any 1 of treatment failure, infectious complications, or grade 4 AE; (3) alive with any 2 of treatment failure, infectious complications, or grade 4 AE; (4) alive with all 3 of treatment failure, infectious complications, or grade 4 AE; and (5) dead.

Patients randomized to trimethoprim-sulfamethoxazole for MRSA BSI had a 44% chance of having a better DOOR compared to those randomized to vancomycin (95% CI, 32% to 55%; Table 4). Distribution of ranking for this study is shown in Figure 3.

Table 2. Desirability of Outcome Ranking Endpoint Based on Clinician Rankings

Rank	Alive	How Many of the Following: Treatment Failure Infectious Complications Ongoing Symptoms Grade 4 Adverse Events
1	Yes	0 of 4
2	Yes	1 of 4
3	Yes	2 of 4
4	Yes	3 of 4
5	Yes	4 of 4
6	No (death)	Any

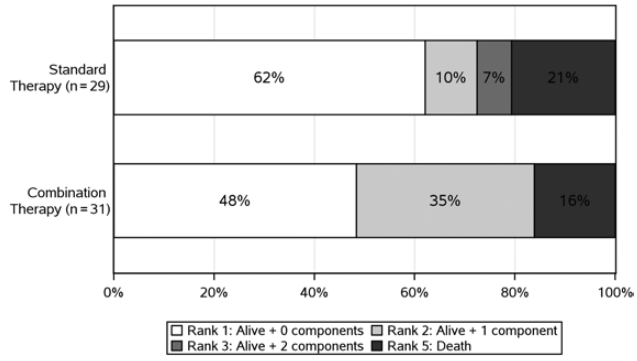

Figure 2. Desirability of outcome ranking distribution for the Combination Antibiotics for Methicillin-Resistant *Staphylococcus aureus*-I trial by treatment group.

Table 3. Combination Antibiotics for Methicillin Resistant *Staphylococcus aureus*-I Trial Desirability of Outcome Ranking Endpoint With Distribution by Treatment Group [10]

Combination Antibiotics for Methicillin Resistant *Staphylococcus aureus*-I Trial DOOR Endpoint			DOOR Distribution		Partial Credit Score			
Rank	Alive	How Many of the Following: Treatment Failure Infectious Complications Grade 4 Adverse Events	Combination Therapy (n = 31)	Standard Therapy (n = 29)	Scenario A	Scenario B	Scenario C	Scenario D
1	Yes	0 of 3	15 (48%)	18 (62%)	100	100	100	100
2	Yes	1 of 3	11 (35%)	3 (10%)	100	0	100	75
3	Yes	2 of 3	0 (0%)	2 (7%)	100	0	0	50
4	Yes	3 of 3	0 (0%)	0 (0%)	100	0	0	25
5	No (death)	Any	5 (16%)	6 (21%)	0	0	0	0

	DOOR Results	Partial Credit Results			
	Probability (95% CI) of a Higher DOOR Using Combination Therapy vs Standard Therapy	Estimated Between-treatment Partial Credit Difference[a,b] (95% CI)			
		A	B	C	D
	0.47 (0.33, 0.60)	4.6 (–16, 25)	–13.7 (–40, 12)	11.5 (–10, 33)	1.7 (–18, 21)
		P = .74	*P* = .31	*P* = .35	*P* = .62

DOOR probability and partial credit scenarios are presented.

Abbreviations: CI, confidence interval; DOOR, desirability of outcome ranking.

[a] Estimated between-treatment difference is equal to mean combination therapy score minus mean standard therapy score.

[b] Exact Wilcoxon *P* values are presented for scenarios A through D. *P* values using *t* test ranged from 0.29 to 0.86.

Table 4. Paul et al [11] Desirability of Outcome Ranking Endpoint With Distribution by Treatment Group

Paul et al [11] DOOR Endpoint			DOOR Distribution		Partial Credit Score			
Rank	Alive	How Many of the Following: Treatment Failure Infectious Complications Grade 4 Adverse Events	TMP/SMX (n = 41)	Vancomycin (n = 50)	Scenario A	Scenario B	Scenario C	Scenario D
1	Yes	0 of 3	16 (39%)	21 (42%)	100	100	100	100
2	Yes	1 of 3	9 (22%)	18 (36%)	100	0	100	75
3	Yes	2 of 3	2 (5%)	1 (2%)	100	0	0	50
4	Yes	3 of 3	0 (0%)	1 (2%)	100	0	0	25
5	No (death)	Any	14 (34%)	9 (18%)	0	0	0	0

	DOOR Results	Partial Credit Results			
	Probability (95% CI) of a Higher DOOR Using TMP/SMX Therapy vs Vancomycin Therapy	Estimated Between-treatment Partial Credit Difference[a,b] (95% CI)			
		A	B	C	D
	0.44 (0.32, 0.55)	–16.2 (–34, 2)	–3.0 (–24, 18)	–17.0 (–36, 2)	–12.6 (–29, 4)
		P = .09	*P* = .83	*P* = .11	*P* = .28

DOOR probability and partial credit scenarios are presented.

Abbreviations: CI, confidence interval; DOOR, desirability of outcome ranking; SMX, sulfamethoxazole; TMP, trimethoprim.

[a] Estimated between-treatment difference is equal to mean TMP/SMX score minus mean vancomycin score.

[b] Exact Wilcoxon *P* values are presented for scenarios A through D. *P* values using *t* test ranged from 0.08 to 0.78.

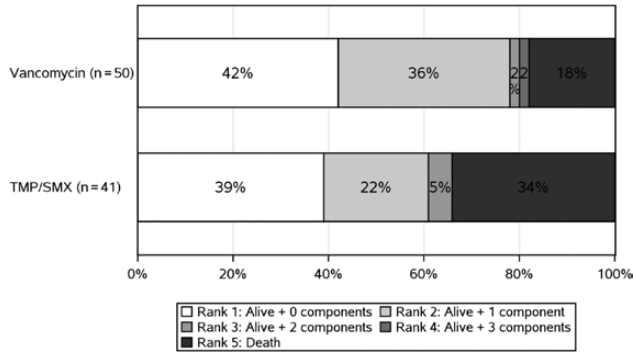

Figure 3. Desirability of outcome ranking distribution for the Paul et al [11] trial by treatment group. Abbreviations: SMX, sulfamethoxazole; TMP, trimethoprim.

As with the CAMERA-1 trial, between-treatment differences were calculated for a range of partial credit scores between 0 and 100 assigned to rankings 2–4 (Figure 3). Partial credit scores were ranged such that 0 ≤ rank 4 ≤ rank 3 ≤ rank 2 ≤ 100. Supplementary Table 1 and Figure 2 display the full range of potential partial credit assignments with several illustrative scenarios. There were no statistically significant differences between treatment arms by any assigned partial credit strategy.

DISCUSSION

Using *S. aureus* BSI as a prototype, we have provided an example of an approach to development and application of a DOOR endpoint that can be utilized in clinical trials to inform pragmatic clinical decision-making. The DOOR outcome generated from our survey of expert ID clinicians may be viewed as a comprehensive evaluation of the global patient experience as it represents a composite of efficacy and safety data. Evaluation of this outcome may thus be considered a structured benefit-to-risk assessment with broader utility for treatment decision-making than traditional methods, which require less intuitive analysis of a series of different proportions, ratios, or means for each outcome of interest (ie, mortality, AEs, quality of life). When comparing *S. aureus* BSI patient profiles, respondents in our survey placed value not just on cure but also on resolution of symptoms and avoidance of AEs. These priorities have been incorporated into our DOOR endpoint that, in contrast to traditional clinical trial outcomes, allows for assessment of the complex nuances of component treatment outcomes and incorporation of competing risks. Respondents were consistent in ranking cases with good outcomes favorably, which would be expected as any outcome measure can distinguish the extremes. Where DOOR analysis stands out most, though, is in the cases with multiple mild to severe complications or adverse outcomes. It is the patient with an eventful course whom the DOOR endpoint will differentiate more than a traditional binary clinical outcome of cure or no cure, or alive or dead. Furthermore, the survey and ensuing DOOR helps to establish which outcomes are meaningfully different and which are interchangeable. Beyond this, the partial credit strategy allows for personalized decision-making as a clinician and patient sharing a decision can weigh specific DOOR strata based on specific priorities.

We applied modification of our constructed DOOR endpoint retrospectively to 2 previously completed trials. The results of the DOOR analysis align with the original clinical trial endpoints for the CAMERA-1 trial, which is not surprising given the negative results of this trial and the relatively small sample size. Though the primary study found a trend toward shortened bacteremia duration with addition of flucloxacillin to vancomycin, the negative results of the DOOR endpoint analysis highlight the fact that a 1-day difference in duration of bacteremia may not be clinically relevant. Though Davis and colleagues [10] took several steps to address competing risks in the original study, DOOR intuitively addresses this issue by incorporating all important outcomes into one composite measure. For example, one would presume that shorter duration of bacteremia is preferred, although patients who die early will also have short durations of bacteremia. In order to address this competing risk, imputation and/or censoring must be done with traditional analyses. However, these approaches remain flawed. If one censors at the time of death, then dying sooner will actually reduce the mean duration of bacteremia. If another number is imputed, such as the end of the study, this is an arbitrary choice and makes the interpretation of the value of the mean number difficult and contingent only on the context of survival.

The DOOR analysis of the Paul trial [11] also echoes the original study results, with the caveat that the initial study included populations other than those with BSI. As with the analysis described above, the original trial found nonstatistically significant trends toward worse outcomes for treatment of BSI with trimethoprim-sulfamethoxazole compared to treatment with vancomycin, including numerical trends toward treatment failure, mortality, and bacteriological failure. However, all of the CIs for the effect estimates crossed zero, suggesting that this study may have been underpowered for the BSI subgroup.

Strengths of our study included involvement in the survey of informed expert clinicians who routinely care for patients. The resultant novel DOOR outcome incorporates the respondents' desire to give weight to ongoing symptoms of infection and AEs, in addition to cure and other infection-related outcomes, and recognizes the cumulative aspect of these events. The final DOOR product includes a comprehensive, valuable, patient-centered benefit-to-risk evaluation and will be a useful tool for future studies. Sample size calculation for a trial using a DOOR endpoint involves testing the null hypothesis that a patient randomly assigned to an experimental strategy will have an X% chance of a better DOOR than one assigned to the control strategy; 50% is often selected for X, noting that >50% implies superiority of the experimental strategy. Using this paradigm, sample size can be calculated with standard software using the Mann-Whitney *U*

test. To test the null hypothesis that the treatment group means are equal when implementing a prespecified partial credit scoring strategy, treatment-specific means and standard deviations based on assumed DOOR distributions for the experimental and control groups can be obtained for sample size determination. Incorporating the partial credit strategy addresses the concern about composite endpoints that differences in specific important component outcomes (eg, death) may be hidden by the composite nature of DOOR [15, 16]. The partial credit strategy directly deals with the concern of unequal importance of component outcomes by allowing special weight placed on more important components such as mortality [4]. The partial credit strategy also gives providers and patients the freedom to choose a treatment based on how they value the outcome. However, for trials, prespecification and transparency are important factors, so one approach to handle this would be to choose a mean partial credit strategy that would serve as the primary scoring strategy. The concern about component outcomes being hidden by the composite nature of the DOOR can also be addressed through evaluation of individual components of particular interest, including mortality, as is routinely done with other composite outcomes [17]. If desired, trials could be sized to compare important component outcomes (eg, survival) in addition to analyses of the DOOR.

We highlight limitations to our study and areas for future investigation. First, we did not elicit patient perspective for these cases. Future iterations of a *S. aureus* BSI DOOR endpoint should incorporate validated measures of quality of life for this population, which may ultimately take the place of the measure of ongoing symptoms. This may help address the fact that certain outcomes within a DOOR stratum have more impact on patient function than others. Second, the survey respondents were selected from the ARLG, which is an active clinical research group that may have differing assessments of the patient experience compared with ID clinicians not involved in research, non-ID–trained physicians, and other providers. Third, the case vignettes simplified actual patient experience and therefore may omit components of the patient experience that would be of value in determining overall desirability of outcome. For example, follow-up duration of 8 weeks was shorter than might be used in a trial. In future trials, we recommend incorporating DOOR endpoint analyses at set follow-up times from randomization to avoid the bias that comes from different follow-up times if based on the end of therapy. Likewise, judgment calls were made around defining surrogate markers for bad outcomes from the patient perspective. For example, persistent bacteremia was a defined variable meant to indicate patients who were failing therapy. However, patient-centered quality-of-life metrics may better assess for bad outcomes since bacteremia duration is only important in the context of impact on the patient. Prospectively designed studies that incorporate the DOOR endpoint will be important next steps given that we had to modify the DOOR endpoint to accommodate the data that had already been collected in these trials, which particularly limited the ability to evaluate ongoing symptoms. As illustrated by this survey and the secondary trial analysis, development of an ordinal outcome is necessarily complex in order to better reflect real-world patient outcomes in complex diseases such as *S. aureus* BSI.

In summary, we have described the development and testing of a novel clinician-driven DOOR endpoint for *S. aureus* BSI that could ultimately supplement future clinical trial design and allow for a global assessment of patient experience. This approach can be used as a template for development of DOOR endpoints for other disease states.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. We are grateful to Natalia Gouskova for assistance with initial statistical analyses.

Financial support. This study was supported in part by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH; award UM1AI104681). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. V. G. F. was supported by the NIH (K24-AI093969). J. S. D. and S. Y. T. received salary support from Australia's National Health and Medical Research Council (Career Development Fellowships 1083105 and 1145033, respectively).

Potential conflicts of interest. S. B. D. reports personal fees from Genentech and personal fees from Actelion outside the submitted work. S. Y. C. T. reports grants from Australian National Health and Medical Research Council during the conduct of the study and personal fees from Genentech outside the submitted work. H. W. B. reports personal fees from Merck outside the submitted work. S. E. C. reports personal fees from Novartis and personal fees from Theravance outside the submitted work. S. R. E. reports grants from National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) during the conduct of the study and personal fees from Takeda/Millennium, Pfizer, Roche, Novartis, Achaogen, Huntington's Study Group, Auspex, Alcon, Merck, Chelsea, Mannkind, QRx Pharma, Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks, Genentech, Affymax, FzioMed, Amgen, GlaxoSmithKline, Boehringer-Ingelheim, American Statistical Association, US Food and Drug Administration, Osaka University, City of Hope, National Cerebral and Cardiovascular Center of Japan, NIH, Muscle Study Group, Society for Clinical Trials, Drug Information Association, University of Rhode Island, New Jersey Medical School/Rutgers, Preclinical Pain Research Consortium for Investigating Safety and Efficacy (PPRECISE) Working Group, Statistical Communications in Infectious Diseases, Cubist, AstraZeneca, Teva, Repros, Austrian Breast & Colorectal Cancer Study Group/Breast International Group, Alliance Foundation Trials, Zeiss, Dexcom, American Society for Microbiology, Taylor and Francis, Claret Medical, Vir, Arrevus, Five Prime, Shire, Alexion, Gilead, Spark, Clinical Trials Transformation Initiative, Nuvelution, and Tracon outside the submitted work. V. G. F. reports grants from NIH, Centers for Disease Control and Prevention, MedImmune, Cerexa/Forest/Actavis/Allergan, Pfizer, Advanced Liquid Logics, Theravance, Novartis, Cubist/Merck, Medical Biosurfaces, Locus, Affinergy, Contrafect, Karius, and Genentech; consultancies with Pfizer, Novartis, Galderma, Novadigm, Durata, Debiopharm, Genentech, Achaogen, Affinium, Medicines Co., Cerexa, Tetraphase, Trius, MedImmune, Bayer, Theravance, Cubist, Basilea, Affinergy, Janssen, xBiotech, and Contrafect; personal fees from Green Cross, Cubist, Cerexa,

Durata, Theravance, and Debiopharm; and royalties from UpToDate and has a patent Sepsis Diagnostics pending. T. L. H. reports grants from NIH during the conduct of the study and personal fees from Motif Bio, Basilea Pharmaceutica, Genentech, and Theravance outside the submitted work. All remaining authors: no reported 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. Evans SR, Follmann D. Comment: fundamentals and innovation in antibiotic trials. Stat Biopharm Res **2015**; 7:331–6.
2. Evans SR, Rubin D, Follmann D, et al. Desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR). Clin Infect Dis **2015**; 61:800–6.
3. Mauri L, D'Agostino RB Sr. Challenges in the design and interpretation of noninferiority trials. N Engl J Med **2017**; 377:1357–67.
4. Evans SR, Follmann D. Using outcomes to analyze patients rather than patients to analyze outcomes: a step toward pragmatism in benefit:risk evaluation. Stat Biopharm Res **2016**; 8:386–93.
5. Holland TL, Arnold C, Fowler VG Jr. Clinical management of *Staphylococcus aureus* bacteremia: a review. JAMA **2014**; 312:1330–41.
6. Fowler VG Jr, Boucher HW, Corey GR, et al.; *S. aureus* Endocarditis and Bacteremia Study Group. Daptomycin versus standard therapy for bacteremia and endocarditis caused by *Staphylococcus aureus*. N Engl J Med **2006**; 355:653–65.
7. Boucher HW. Challenges in anti-infective development in the era of bad bugs, no drugs: a regulatory perspective using the example of bloodstream infection as an indication. Clin Infect Dis **2010**; 50 Suppl 1:S4–9.
8. Thwaites GE, Scarborough M, Szubert A, et al.; United Kingdom Clinical Infection Research Group. Adjunctive rifampicin for *Staphylococcus aureus* bacteraemia (ARREST): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet **2018**; 391:668–78.
9. Keller SC, Williams D, Gavgani M, et al. Rates of and risk factors for adverse drug events in outpatient parenteral antimicrobial therapy. Clin Infect Dis **2018**; 66:11–9.
10. Davis JS, Sud A, O'Sullivan MVN, et al. Combination of vancomycin and beta-lactam therapy for methicillin-resistant *Staphylococcus aureus* bacteremia: A pilot multicenter randomized controlled trial. Clin Infect Dis **2016**; 62: 173–80.
11. Paul M, Bishara J, Yahav D, et al. Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by meticillin resistant *Staphylococcus aureus*: randomised controlled trial. BMJ **2015**; 350.
12. Guidance for industry and investigators: Safety reporting requirements for INDs and BA/BE studies, p. 1–29. Available at: https://www.fda.gov/downloads/Drugs/Guidances/UCM227351.pdf. Accessed 28 February 2018.
13. Common terminology criteria for adverse events (CTCAE). Available at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_50. Accessed 28 February 2018.
14. Tong SY, Nelson J, Paterson DL, et al. CAMERA2 - combination antibiotic therapy for methicillin-resistant *Staphylococcus aureus* infection: study protocol for a randomised controlled trial. Trials **2016**; 17:170.
15. Phillips PP, Morris TP, Walker AS. DOOR/RADAR: a gateway into the unknown? Clin Infect Dis **2016**; 62:814–5.
16. Evans S, Follmann D, Schoenfeld D, Fowler VG Jr, Chambers HF. Reply to Phillips, Morris, and Walker. Clin Infect Dis **2016**; 62:815–6.
17. Multiple endpoints in clinical trials guidance for industry (draft guidance). Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm536750.pdf. Accessed 28 February 2018.