

Improving Traditional Registrational Trial End Points: Development and Application of a Desirability of Outcome Ranking End Point for Complicated Urinary Tract Infection Clinical Trials

Jessica Howard-Anderson,^{1,*} Toshimitsu Hamasaki,² Weixiao Dai,² Deborah Collyar,³ Daniel Rubin,⁴ Sumathi Nambiar,⁵ Tori Kinamon,⁴ Carol Hill,⁶ Steven P. Gelone,⁷ David Mariano,⁷ Takamichi Baba,⁸ Thomas L. Holland,^{6,9} Sarah B. Doernberg,¹⁰ Henry F. Chambers,¹⁰ Vance G. Fowler Jr,^{6,9} Scott R. Evans,² and Helen W. Boucher¹¹; on behalf of the Antibacterial Resistance Leadership Group

¹Department of Medicine, Division of Infectious Diseases, Emory University School of Medicine, Atlanta, Georgia, USA; ²Biostatistics Center and Department of Biostatistics and Bioinformatics, Milken Institute School of Public Health, George Washington University, Washington, DC, USA; ³Patient Advocates in Research, Danville, California, USA; ⁴Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; ⁵Johnson & Johnson, Raritan, New Jersey, USA; ⁶Duke Clinical Research Institute, Durham, North Carolina, USA; ⁷Nabriva Therapeutics US, Inc, Fort Washington, Pennsylvania, USA; ⁸Biostatistics Center, Shionogi & Co, Ltd, Osaka, Japan; ⁹Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA; ¹⁰Department of Medicine, Division of Infectious Diseases, University of California, San Francisco, California, USA; and ¹¹Tufts University School of Medicine and Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Boston, Massachusetts, USA

(See the Editorial Commentary by Rodríguez-Baño and Gutiérrez-Gutiérrez on pages 1175–6.)

Background. Traditional end points used in registrational randomized, controlled trials (RCTs) often do not allow for complete interpretation of the full range of potential clinical outcomes. Desirability of outcome ranking (DOOR) is an approach to the design and analysis of clinical trials that incorporates benefits and risks of novel treatment strategies and provides a global assessment of patient experience.

Methods. Through a multidisciplinary committee of experts in infectious diseases, clinical trial design, drug regulation, and patient experience, we developed a DOOR end point for infectious disease syndromes and demonstrated how this could be applied to 3 registrational drug trials (ZEUS, APEKS-cUTI, and DORI-05) for complicated urinary tract infections (cUTIs). ZEUS compared fosfomycin to piperacillin/tazobactam, APEKS-cUTI compared cefiderocol to imipenem, and DORI-05 compared doripenem to levofloxacin. Using DOOR, we estimated the probability of a more desirable outcome with each investigational antibacterial drug.

Results. In each RCT, the DOOR distribution was similar and the probability that a patient in the investigational arm would have a more desirable outcome than a patient in the control arm had a 95% confidence interval containing 50%, indicating no significant difference between treatment arms. DOOR facilitated improved understanding of potential trade-offs between clinical efficacy and safety. Partial credit and subgroup analyses also highlight unique attributes of DOOR.

Conclusions. DOOR can effectively be used in registrational cUTI trials. The DOOR end point presented here can be adapted for other infectious disease syndromes and prospectively incorporated into future clinical trials.

Keywords. desirability of outcome ranking; complicated urinary tract infections; clinical trials; drug development.

With increasing antibiotic resistance and a dwindling antibacterial pipeline, clinical trials to evaluate novel anti-infective agents are critical [1, 2]. For antibacterial registrational studies, we typically rely on noninferiority, phase 3 clinical trials to evaluate the safety and efficacy of new antibiotics and to support regulatory approval. While these trials provide meaningful information that can allow for drug approval, the results often

do not directly inform which therapy is best for an individual patient [3–5]. Typical trial end points are binary, failing to distinguish important finer gradations of clinical responses that can occur over the course of treatment [3–5]. Additionally, efficacy and safety outcomes are analyzed separately, making it difficult to combine these results into a single assessment to inform treatment selection.

Regulatory agencies, including the US Food and Drug Administration (FDA), have encouraged pragmatic study designs and improved outcome assessments in registrational drug trials [6–8]. The Antibacterial Resistance Leadership Group (ARLG) has developed and pioneered the use of desirability of outcome ranking (DOOR), an approach to the design and analyses of clinical trials that uses an ordinal ranking to assess clinical outcomes. DOOR incorporates both the benefits

Received 27 May 2022; editorial decision 11 August 2022; published online 29 August 2022

Correspondence: J. Howard-Anderson, Emory University School of Medicine 3, Orr Building #1016, 550 Peachtree St NE, Atlanta, GA 30303 (jrhwa4@emory.edu).

Clinical Infectious Diseases® 2023;76(3):e1157–e65

© The Author(s) 2022. Published by Oxford University Press on behalf of Infectious Diseases Society of America. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

https://doi.org/10.1093/cid/ciac692

and risks of an investigational treatment strategy into a single outcome, providing a global assessment of a patient's experience [3, 4]. ARLG has applied DOOR in observational studies [9–11] and has designed prospective trials with DOOR as the primary end point (ClinicalTrials.gov: NCT04775953 and NCT02891915) [12]. Extending this work, ARLG created a multidisciplinary committee comprised of experts in infectious diseases, clinical trial design, drug regulation, and patient advocacy, including members from academia, FDA, the National Institutes of Health, and the pharmaceutical industry, to create a standardized DOOR end point that could be used in phase 2/3 registrational drug trials.

Here, we used data from 3 previously published, noninferiority, double-blind, randomized, controlled trials (RCTs) that addressed complicated urinary tract infections (cUTIs) [13–15]. Our primary objectives were to develop an infectious diseases DOOR that could be used in cUTI trials, demonstrate that this approach can be implemented in antibacterial registrational trials, and explain how DOOR analyses can be presented and interpreted.

METHODS

Development of Door Analysis Strategy

In May 2020, the ARLG created an Innovations Working Group that included experts in clinical trials, anti-infective drug development and regulations, statistical analysis, quality of life, and patient advocacy [16]. This committee developed and agreed on a generalized DOOR analysis strategy that was adapted from prior work in *Staphylococcus aureus* bloodstream infections [5] and could be applied to clinical trials for common infectious disease syndromes (Figure 1A). Through iterative feedback and consensus building, the committee agreed on how each DOOR event would be defined for cUTIs (Figure 1B), as well as for other infectious syndromes including acute bacterial skin and skin structure infections (ABSSSI), hospital-acquired or ventilator-associated bacterial pneumonia (HABP/VABP), and complicated intraabdominal infections (cIAI) (not shown).

Overview and Design of the Primary cUTI Studies

We approached 6 pharmaceutical sponsors of registrational trials for cUTIs. Two sponsors agreed to share deidentified data, contributed in-kind from 3 trials (ZEUS, APEKS-cUTI, and DORI-05) [13–15]. These studies were noninferiority, double-blind, RCTs of patients with cUTIs; we retrospectively analyzed them using the described DOOR end point. These studies have been previously published, and key characteristics are presented in Table 1. ZEUS compared intravenous (IV) fosfomycin to IV piperacillin/tazobactam [13]. APEKS-cUTI compared IV cefiderocol to IV imipenem [14]. DORI-05 compared IV doripenem to IV levofloxacin, with a step-down option of oral levofloxacin in both groups after 3 days of IV therapy [15].

DORI-05, the oldest cUTI study analyzed, completed enrollment in 2006 and used microbiologic eradication as the primary end point without including a clinical component that is now recommended by FDA guidance [15, 17]. DORI-05 collected data on clinical cure as a secondary end point, which allowed us to incorporate this into our DOOR analysis. All trials used a similar time frame for the test of cure (TOC) visit (range, day 14–25); however, APEKS-cUTI and DORI-05 monitored adverse events (AEs) for a longer duration than ZEUS. For DORI-05, we noted a discrepancy between the TOC time frame stated in the protocol (end of therapy plus 6–9 days) and the published article (end of therapy plus 5–11 days) [15]. For our primary DOOR analysis, we used the TOC time frame stated in the published article (Table 1); however, as a sensitivity analysis, we performed the DOOR analysis using the time frame stated in the protocol, and the results were similar (data not shown).

Application of DOOR Analysis Strategy and Abstraction of Key Variables

Using the DOOR analysis strategy (Figure 1), we retrospectively assigned each study participant a mutually exclusive rank of 1 through 5. Rank 1 represented the most desirable outcome and included anyone who was alive and did not experience any of the undesirable, prespecified events. Rank 5 represented the least desirable outcome and included all patients who died. Ranks 2 through 4 included patients who were alive but had 1, 2, or 3 events, respectively. The events included in the DOOR analysis were categorized as absence of clinical response, infectious complications, and serious AEs (SAEs; Figure 1).

We abstracted data on “absence of clinical response” from each study's clinical response variable performed at the TOC. While each study defined clinical response slightly differently, all studies required patients to have resolution or improvement of cUTI symptoms without recurrence of symptoms at TOC. We did not include a microbiologic eradication component. Patients who did not meet the study's clinical cure definition (ie, patients with clinical failure, indeterminate or missing outcomes) were defined as having an absence of clinical response in the DOOR analysis. In sensitivity analyses, we modified the adjudication of patients with indeterminate or missing outcomes in the following 3 ways: patients with indeterminate/missing outcomes were ranked above those with clinical failure when they otherwise had the same rank (“tie-breaker” analysis), patients with indeterminate/missing outcomes were counted as having clinical cure, and patients with indeterminate/missing outcomes were excluded.

For the DOOR category of infectious complications, 2 board-certified infectious diseases physicians (J. H. A. and H. W. B.) reviewed all recorded AEs from the trials and determined if any event met criteria for 1 of the predefined cUTI infectious complications (renal or intraabdominal abscess, septic shock, bacteremia with the same bacteria identified in the original

A Generalized DOOR Analysis Strategy

Rank ^a	Alive	How many of following events: 1. Absence of clinical response ^b 2. Infectious complications ^c 3. Serious adverse events ^d
1 (most desirable)	Yes	0 of 3
2	Yes	1 of 3
3	Yes	2 of 3
4	Yes	3 of 3
5 (least desirable)	No (death)	Any

B Definitions for cUTI Trials

Event Category	ARLG Criteria for cUTI Trials
Absence of clinical response ^b	<ul style="list-style-type: none"> • Did not meet clinical success or cure as assessed by study investigator at test of cure • Recurrent cUTI prior to test of cure
Infectious complications ^c	<ul style="list-style-type: none"> • Renal or intraabdominal abscess • Septic shock • Bacteremia due to the same bacteria identified in original urine culture • Recurrent UTI or pyelonephritis after test of cure • <i>Clostridioides difficile</i> • Epididymo-orchitis^e • Prostatic abscess^e
Serious adverse events ^d	<ul style="list-style-type: none"> • Any untoward medical event that: <ul style="list-style-type: none"> ○ Results in death ○ Is life-threatening ○ Requires inpatient hospitalization or prolongation of existing hospitalization ○ Results in persistent or significant disability/incapacity or ○ Is a congenital anomaly/birth defect

Figure 1. DOOR analysis strategy. *A*, The generalized DOOR analysis strategy that could be applied to any infectious diseases clinical trial. *B*, Details of how the DOOR component events were defined a priori for cUTI trials. Abbreviations: ARLG, Antibacterial Resistance Leadership Group; cUTI, complicated urinary tract infection; DOOR, desirability of outcome ranking; UTI, urinary tract infection. ^aQuality-of-life markers, when available, could be used as a tiebreaker for patients with the same rank. ^bDefined as lack of global resolution of index infection or recurrence of index infection before test of cure. ^cDefined as a newly identified complication or progression of the original infection that was not present at enrollment, including the development of *Clostridioides difficile*. ^dDefined according to International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) E6 Good Clinical Practice guidelines. ^eAdded after the initial review of adverse events from the cUTI trials with agreement by the ARLG Innovations Committee.

urine culture, recurrent UTI or pyelonephritis after TOC, or *Clostridioides difficile*). The physicians were blinded to treatment assignment and followed a standard operating procedure reviewed by all committee members. Agreement between the reviewers had to be unanimous, and any events that were unable to be resolved were taken back to the full committee. During the review, 2 additional AE terms (epididymo-orchitis and prostatic abscess) were included as infectious complications as they were determined by the committee to be consistent with the definition. All infectious complications had to be identified after patient enrollment and occur during the period each trial set for AE monitoring (Table 1).

SAEs were defined according to the International Council for Harmonisation of Technical Requirements for

Pharmaceuticals for Human Use (ICH) E6 Good Clinical Practice guidelines, which are endorsed by the FDA (21 CFR 312.32) [18]. An AE was coded as serious if the event resulted in death, was life-threatening, required inpatient hospitalization or prolongation of existing hospitalization, resulted in persistent or significant disability/incapacity, or was a congenital anomaly/birth defect. We included all SAEs and deaths that occurred during the same follow-up time that AEs were monitored for (Table 1).

Statistical Analyses

Our primary analysis used the modified intention-to-treat (mITT) population for each study, defined as all randomized patients who received at least 1 dose of the study drug. This

Table 1. Study Characteristics of the 3 Randomized Controlled Trials Analyzed Using Desirability of Outcome Ranking

Characteristic	ZEUS	APEKS-cUTI	DORI-05
Number of participants, intention to treat (modified intention to treat) ^a	465 (464)	452 (448)	753 (748)
Study design	Multicenter, double-blind, noninferiority, RCT	Multicenter, double-blind, noninferiority, RCT	Multicenter, double-blind, noninferiority, RCT
Years of enrollment	May 2016–January 2017	February 2015–August 2016	December 2003–March 2006
Study drugs	Fosfomycin vs piperacillin/tazobactam	Cefiderocol vs imipenem	Doripenem vs levofloxacin
Primary diagnosis, n (%) ^b			
Complicated urinary tract infection	223 (48)	333 (74)	427 (56)
Acute pyelonephritis	242 (52)	119 (26) ^c	326 (43) ^c
Duration of therapy	7 days (14 if bacteremia)	7–14 days	10 days (14 days if bacteremia) ^d
TOC timeframe	Day 19 (+2)	EOT + 7 days (±2) (approximately day 14–21)	EOT + 5–11 days ^e (approximately day 15–25)
Time frame for monitoring adverse events	Day 26 (±2)	EOT + 28 days (approximately day 35–42)	EOT + 28–42 days (approximately day 38–56)
Original primary outcome	Overall success (clinical cure and microbiologic eradication) at TOC	Overall success (clinical cure and microbiologic eradication) at TOC	Microbiologic eradication at TOC

Abbreviations: EOT, end of study drug therapy; RCT, randomized, controlled trial; TOC, test of cure.

^aModified intention to treat was defined as all randomized patients who received any amount of study drug.

^bBased on intention-to-treat population.

^cDefined as acute, uncomplicated pyelonephritis.

^dPatients in both treatment groups were eligible to switch to oral levofloxacin after 3 days of intravenous therapy if prespecified clinical criteria were met.

^eAs published in Naber et al [15]. The original study protocol stated EOT + 6–9 days.

was different than the population used in the primary studies' analyses, which included a microbiologic criterion. For each study, we compared the DOOR distribution between treatment groups and calculated the probability of a more desirable outcome with one treatment compared to the other (DOOR probability; Wilcoxon Mann–Whitney *U* statistic adjusted for tie) with a corresponding 2-sided 95% confidence interval (CI) [19]. We also calculated this probability for each DOOR component. A DOOR probability of 50% indicates no difference.

Additionally, we defined and analyzed prioritized DOORs, one prioritizing efficacy and another prioritizing safety. When comparing 2 patients with the same number of events, the DOOR prioritizing efficacy prioritized avoidance of clinical failure over SAEs or infectious complications and the DOOR prioritizing safety prioritized avoidance of SAEs and infectious complications over clinical failure.

We performed subgroup analyses that were decided by the committee prior to data analysis to assess treatment effect heterogeneity based on patient characteristics, including age, infection type (cUTI or acute pyelonephritis), and creatinine clearance. The DOOR component and subgroup analyses were not adjusted for multiple comparisons.

Partial Credit Analysis

We performed a partial credit analysis using 3 hypothetical scoring keys (scenarios A, B, and C) [3]. In this analysis, the DOOR categories are scored like an academic test. The most desirable rank (rank 1) is assigned a score of 100 and the least desirable rank (rank 5) is assigned a 0. Ranks 2–4 are assigned

“partial credit” (a score between 0 and 100) while maintaining the original rank order [3]. Patients or clinicians may customize the partial credit analysis by selecting a grading key based on their own preferences. For example, one patient may value survival over all other outcomes, whereas another may only want to survive if they are able to maintain a certain lifestyle. Partial credit scoring can be prespecified in clinical trial design for transparency, informed from quality-of-life studies or from a survey of expert clinicians regarding a grading key. Treatment comparisons are made by comparing mean partial credit scores between treatments, using the Welch *t* test and related methods. A 95% CI for the difference that contains zero indicates no significant difference in mean partial scores between treatment groups. The advantages of the partial credit scoring approach are that it strategically scores the DOOR categories to account for nonuniform steps between categories and it has an intuitive interpretation given the 100-point scale.

All analyses were conducted using SAS v9.4 (SAS Institute, Cary, NC) or R statistical software R 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

All 3 trials had more than 400 participants included in the mITT analysis (range, 448–748; Table 1). Each study demonstrated that the investigational drug being studied (ie, fosfomycin, cefiderocol, and doripenem) was noninferior to the comparator in the primary analysis. The trials did not identify any significant safety concerns [13–15]. Doripenem and cefiderocol are now FDA approved for the treatment of cUTIs,

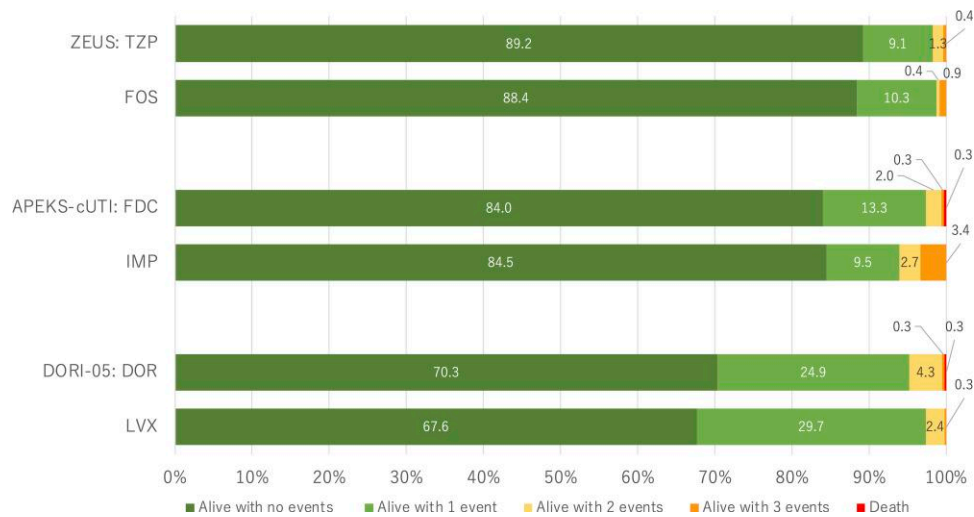


Figure 2. Desirability of outcome ranking distribution by treatment groups for each clinical trial analyzed. The events are defined in Figure 1 and include absence of clinical response, infectious complications, serious adverse events, and death. Abbreviations: DOR, doripenem; FDC, cefiderocol; FOS, fosfomycin; IMP, imipenem-cilastatin; LVX, levofloxacin; TZP, piperacillin/tazobactam.

and fosfomycin's new drug application is pending due to the coronavirus disease 2019 pandemic, awaiting FDA inspections of foreign manufacturing facilities.

The DOOR distribution was not significantly different in each study. Most patients were in rank 1, alive without any undesirable events (range, 68%–89%; Figure 2). In ZEUS, the probability that a patient in the fosfomycin group would have a more desirable outcome than a patient in the piperacillin/tazobactam group was 49.6% (95% CI, 46.7%–52.5%). In APEKS-cUTI, the probability that a patient in the cefiderocol group would have a more desirable outcome than a patient in the imipenem group was 50.1% (95% CI, 46.4%–53.7%). In DORI-05, the probability that a patient in the doripenem arm would have a more desirable outcome than a patient in the levofloxacin arm was 51.0% (95% CI, 47.6%–54.3%). Sensitivity analyses modifying the handling of indeterminate or missing clinical efficacy cases resulted in similar conclusions for ZEUS or APEKS-cUTI. However, for DORI-05, the trial that had the largest proportion of missing or indeterminate values, there was a significant change in the DOOR probability favoring levofloxacin in 2 of the sensitivity permutations (Supplementary Figure 1).

As a composite outcome, the DOOR components (absence of clinical response, infectious complications, SAEs, and death) should be thoroughly analyzed. Results for each trial are displayed in Figure 3 and Supplementary Table 1. In ZEUS, significant differences between treatment arms were not observed for each component. However, in APEKS-cUTI, patients in the cefiderocol arm had a more desirable outcome when infectious complications were analyzed. In DORI-05, patients in the doripenem arm had a more desirable outcome for the clinical

efficacy variable but a less desirable outcome when infectious complications were assessed, illustrating a risk–benefit trade-off. The DOOR probabilities for the DOORs prioritizing efficacy and safety were similar to the original DOOR (Figure 3).

In subgroup analyses, the DOOR probabilities had wide CIs but were overall similar when treatment arms by age, infection type, and creatinine clearance were compared (Supplementary Figures 2–4).

Partial Credit Analysis

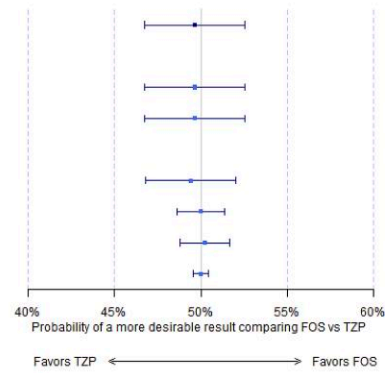
Three hypothetical scenarios that could represent different patient preferences were evaluated using partial credit (Figure 4). In scenario A, the hypothetical patient only cares about surviving the infection (this is equivalent to mortality as the primary end point). For scenario B, the patient places significant value on avoiding complications and only wants to survive the infection if they do not experience any significant undesirable events. The patient in scenario C has a more compromising grading key but still views death as significantly worse than having several, nonfatal AEs. In all 3 trials, we did not observe a significant difference in the mean partial credit scores between treatment arms for any of the patient scenarios (Figure 4).

DISCUSSION

Here, we describe the development of an infectious diseases DOOR end point and demonstrate how this can be applied to cUTI trials. Our DOOR end point includes different and important clinical components, including treatment efficacy, infectious complications, SAEs, and death. These analyses

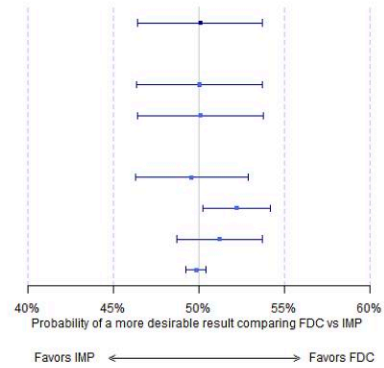
A ZEUS

	TZP (N = 231) n (%)	FOS (N = 233) n (%)	DOOR probability (95% CI)
DOOR			49.6% (46.7%, 52.5%)
Prioritized DOOR			
Prioritized Efficacy			49.6% (46.7%, 52.5%)
Prioritized Safety			49.7% (46.8%, 52.6%)
DOOR components			
Absence of clinical response	19 (8.2%)	22 (9.4%)	49.4% (46.8%, 52.0%)
Infectious complications	5 (2.2%)	5 (2.1%)	50.0% (48.6%, 51.4%)
SAE	6 (2.6%)	5 (2.1%)	50.2% (48.8%, 51.7%)
Death	0 (0.0%)	0 (0.0%)	50.0% (49.6%, 50.4%)



B APEKS-cUTI

	FDC (N = 300) n (%)	IMP (N = 148) n (%)	DOOR probability (95% CI)
DOOR			50.1% (46.4%, 53.7%)
Prioritized DOOR			
Prioritized Efficacy			50.0% (46.4%, 53.7%)
Prioritized Safety			50.1% (46.4%, 53.8%)
DOOR components			
Absence of clinical response	39 (13.0%)	18 (12.2%)	49.6% (46.3%, 52.9%)
Infectious complications	3 (1.0%)	8 (5.4%)	52.2% (50.2%, 54.2%)
SAE	15 (5.0%)	11 (7.4%)	51.2% (48.7%, 53.7%)
Death	1 (0.3%)	0 (0.0%)	49.8% (49.3%, 50.4%)



C DORI-05

	DOR (N = 374) n (%)	LVX (N = 374) n (%)	DOOR probability (95% CI)
DOOR			51.0% (47.6%, 54.3%)
Prioritized DOOR			
Prioritized Efficacy			51.6% (48.3%, 55.0%)
Prioritized Safety			50.3% (46.9%, 53.7%)
DOOR components			
Absence of clinical response	81 (21.7%)	113 (30.2%)	54.3% (51.1%, 57.4%)
Infectious complications	23 (6.1%)	5 (1.3%)	47.6% (46.2%, 49.0%)
SAE	25 (6.7%)	14 (3.7%)	48.5% (46.9%, 50.1%)
Death	1 (0.3%)	0 (0.0%)	49.9% (49.5%, 50.2%)

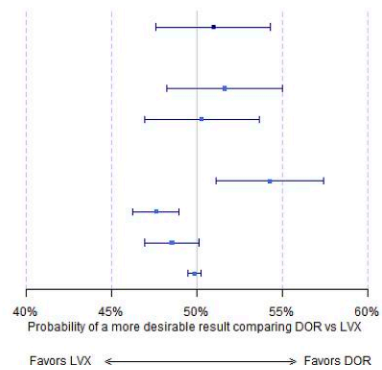


Figure 3. Forest plot demonstrating the DOOR probabilities for the DOOR, DOOR prioritized for efficacy and safety, and the DOOR components (treatment failure, infectious complications, serious adverse events, and death) for each clinical trial (A. ZEUS; B. APEKS-cUTI; C. DORI-05). Abbreviations: CI, confidence interval; DOR, doripenem; DOOR, desirability of outcome ranking; FDC, cefiderocol; FOS, fosfomycin; IMP, imipenem-cilastatin; LVX, levofloxacin; SAE, serious adverse events; TZP, piperacillin/tazobactam.

demonstrate that it is feasible to use DOOR in phase 2/3 multicenter RCTs and that DOOR could be used to evaluate novel anti-infective agents, including as part of the regulatory approval process. The studies analyzed spanned a wide range of years, enrolling participants from 2003 through 2017, supporting the applicability of this analysis. Additionally, DOOR can be used in subgroup analyses that can identify groups of patients with the most favorable risk-to-benefit balance. The

DOOR end point we created can be adapted to fit other infectious diseases and should be prospectively incorporated in future trials.

While the 3 RCTs analyzed in this study were designed as noninferiority studies, the DOOR analysis allows for an assessment of superiority of the investigational drug based on the overall patient experience, the most important question for guiding clinical practice. Overall, the DOOR analysis

A ZEUS

DOOR rank	Scenario A		Scenario B		Scenario C	
1: Alive with no events	100		100		100	
2: Alive with 1 event	100		0		80	
3: Alive with 2 events	100		0		60	
4: Alive with 3 events	100		0		40	
5: Death	0		0		0	
Group means of partial credit scores	FOS	TZP	FOS	TZP	FOS	TZP
	100	100	88.4	89.2	97.3	97.4
Difference in means (95% CI)*	0.0 (NA, NA)		-0.8 (-6.5, 5.0)		-0.1 (-1.7, 1.4)	

* Difference = FOS – TZP

B APEKS-cUTI

DOOR rank	Scenario A		Scenario B		Scenario C	
1: Alive with no events	100		100		100	
2: Alive with 1 event	100		0		80	
3: Alive with 2 events	100		0		60	
4: Alive with 3 events	100		0		40	
5: Death	0		0		0	
Group means of partial credit scores	FDC	IMP	FDC	IMP	FDC	IMP
	99.7	100.0	84.0	84.5	96.0	95.0
Difference in means (95% CI)*	-0.3 (-1.0, 0.3)		-0.5 (-7.7, 6.7)		1.0 (-1.5, 3.5)	

* Difference = FDC – IMP

C DORI-05

DOOR rank	Scenario A		Scenario B		Scenario C	
1: Alive with no events	100		100		100	
2: Alive with 1 event	100		0		80	
3: Alive with 2 events	100		0		60	
4: Alive with 3 events	100		0		40	
5: Death	0		0		0	
Group means of partial credit scores	DOR	LVX	DOR	LVX	DOR	LVX
	99.7	100.0	70.3	67.6	92.9	92.9
Difference in means (95% CI)*	-0.3 (-0.8, 0.3)		2.7 (-4.0, 9.3)		<-0.1 (-1.7, 1.6)	

* Difference = DOR – LVX

Figure 4. DOOR partial credit analysis. Scenario A represents a patient who values only survival (equivalent to a mortality outcome). Scenario B represents a patient who places more value on minimizing complications and would not accept any undesirable event. Scenario C represents a patient who places significant value on survival but balances this with wanting to avoid some complications. For each scenario, the mean of the partial credits scores is calculated for each treatment group and then the difference between groups is obtained. A difference with a 95% CI that overlaps zero indicates no significant difference between groups. Abbreviations: CI, confidence interval; DOR, doripenem; DOOR, desirability of outcome ranking; FDC, cefiderocol; FOS, fosfomycin; IMP, imipenem-cilastatin; LVX, levofloxacin; NA, not applicable; TZP, piperacillin/tazobactam.

demonstrated concordant results to the primary published analysis of each study, that the investigational antibiotic provided comparable efficacy to the established drug [13–15]. However, DOOR provided more detailed information to patients and clinicians about the overall patient outcomes (risks and benefits) and how their experiences compared across treatments. This is most clearly exemplified in DORI-05 where, despite the overall DOOR probability being close to 50%, we observed differences in the components of clinical efficacy and infectious complications. This delineation can help clinicians and patients make more informed decisions on the best treatment regimen and should be part of shared decision making. For example, an elderly patient who is minimally symptomatic with a cUTI but has many comorbidities may want to use levofloxacin over doripenem, which was more likely to result in a desirable outcome when the focus was on infectious

complications and possibly SAEs. A younger patient whose cUTI has caused a significant disruption in their quality of life may prioritize clinical efficacy over AEs and thus prefer doripenem over levofloxacin. This analysis may also be helpful to drug regulators and could potentially be included in the package insert.

In all 3 cUTI studies analyzed, a large majority of patients experienced none of the prespecified undesirable events. However, there are likely differences in this group that we could not ascertain due to limitations in data collection and the inability to differentiate nuances in patient outcomes. Recent work analyzing health-related quality of life (HRQoL) in patients with *S. aureus* and gram-negative bacterial bloodstream infections demonstrated that these infectious conditions have a significant impact on HRQoL [20]. Patient-reported outcomes that involve HRQoL may help further differentiate

patients with similar DOOR ranks. Future clinical trials can be designed to take a proactive and structured approach to systematically capture patient-reported outcomes as part of the DOOR end point, allowing for more complete characterization of the patient experience. Further work is needed to determine how to best incorporate HRQoL into DOOR.

The DOOR framework is strengthened by the fact that it was developed based on prior published work and through a multidisciplinary collaboration of experts. We were able to partner with 2 pharmaceutical companies that did not sponsor this study or participate in the analysis but were willing to share their clinical trial data to further this work. The FDA has previously partnered with the Foundation for National Institutes of Health Biomarkers Consortium [7] and the Clinical Trials Transformation Initiative [8] to explore novel end points for drug trials that target hospital-acquired and ventilator-associated bacterial pneumonia; however, we are unaware of other collaborative efforts that involve such a diverse group of stakeholders in the field of anti-infective drug development.

Our study is limited due to its retrospective approach. We relied on the definition of clinical efficacy derived from FDA guidance and assessed by primary study investigators [17]. We could not collect new data, and this created difficulty in categorizing patients with indeterminate or missing responses. Microbiologic parameters were not included in our DOOR end point as the committee believed that microbiologic persistence, without associated clinical symptoms, did not impact how a patient “feels, functions, or survives.” We focused instead on capturing clinical recurrence of cUTI in the DOOR event categories. The emergence of antimicrobial resistance may be a parameter to include in the future; however, these data were not available. The frequency of infectious complications was abstracted from coded data recorded for all AEs, and we could not ensure standardized reporting of our infectious complications of interest. Last, in registrational trials such as these, high-risk patients are often excluded, which may minimize the number of AEs reported. The infectious complications we included, as well as mortality, occurred infrequently. Future analyses that include a larger number of infectious complications in the DOOR end point may be informative.

In conclusion, this work provides a foundation for using DOOR in clinical trials that investigate cUTIs and other infectious diseases. Prospective studies would benefit from creation of specific case report forms for measuring DOOR based on relevant clinical outcomes, prespecified infectious complications, and important AEs. Future work will involve refining the DOOR end point for cUTIs once it has been used prospectively as well as applying our DOOR end point to other common entry indications for anti-infective agents including ABSSSI, HABP/VABP, and cIAI. DOOR provides crucial information on the patient experience based on the benefits and harms associated with novel anti-infective agents and gives clinicians

and patients more actionable information than is typically gained through current registrational trials.

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. The authors thank all study investigators, staff, and participants who participated in the trials that were analyzed (ZEUS, APEKS-cUTI, and DORI-05).

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH) and should not be construed to represent US Food and Drug Administration (FDA) views or policies.

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the NIH (award UM1AI104681 reported by C. H., V. G. F., S. R. E., and H. F. C.; payments made to institution during the conduct of this analysis). J. H. A. was in part supported by the Antibacterial Resistance Leadership Group Fellowship (NIAID UM1AI104681). This work was also supported in part by an appointment to the Research Participation Program at the FDA administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the FDA reported by T. K.

Potential conflicts of interest. H. W. B. declares consulting fees from Antimicrobial Agents and Chemotherapy, Sanford Guide, and ID Clinics of North America. T. B. is an employee of and holds stock in Shionogi & Co and reports a nondisclosure agreement between Duke University and Shionogi (to share the analysis conducted by Shionogi with the Antibacterial Resistance Leadership Group [ARLG]). H. F. C. reports grant funding from NIH; consulting fees from Johnson & Johnson; participation on a data and safety monitoring board (DSMB) for Merck; and stock in Merck and Moderna. S. B. D. reports research funding from Gilead, Regeneron, and NIH; payments for clinical event adjudication committees from Basilea and Shionogi (pending); and consulting fees from Genentech. S. R. E. reports grant funding from NIH and Degruyter; royalties from Taylor & Francis; consulting fees from Genentech, AstraZeneca, Takeda, Microbiotix, Johnson & Johnson, Endologix, ChemoCentryx, Becton Dickinson, Atricure, Roivant, Neovasc, Nobel Pharma, Horizon, International Drug Development Institute, and SVB Leerink; payments for lectures from the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks; support for attending meetings from FDA, Clinical Trial Transformation Initiative, Council for International Organization of Medical Sciences, International Chinese Statistical Association, participation on DSMBs for NIH, Breast International Group, University of Pennsylvania, Duke University, Roche, Pfizer, Takeda, Akouos, Appelis, Teva, Vir, DayOneBio, Alexion, Tracoon, Rakuten, AbbVie, Nuvelution, Clover, FHI Clinical, Lung Biotech, SAB Biopharm, and Advantagene; and being a board member for the American Statistical Association, Society for Clinical Trials, and Frontier Science Foundation. V. G. F. reports personal fees from Novartis, Debiopharm, Genentech, Achaogen, Affinium, Medicines Co, MedImmune, Bayer, Basilea, Affinergy, Janssen, Contrafect, Regeneron, Destiny, Amphiphilic Biosciences, Integrated Biotherapeutics, C3J, Armata, Valanbio, Akagera, Aridis, Roche; grants from NIH, MedImmune, Allergan, Pfizer, Advanced Liquid Logics, Theravance, Novartis, Merck, Medical Biosurfaces, Locust, Affinergy, Contrafect, Karius, Genentech, Regeneron, Deep Blue, Basilea, Janssen; royalties from UpToDate; stock options from Valanbio and ArcBio; honoraria from the Infectious Diseases of America for service as associate editor of *Clinical Infectious Diseases*; and a patent for sepsis diagnostics is pending. S. P. G. is an employee of Nabriva Therapeutics and holds stock in Nabriva Therapeutics. T. H. reports consulting fees from Tanabe-Mitsubishi Pharma and payment for lectures from Johnson & Johnson and Duke

University. T. L. H. declares royalties from UpToDate and consulting fees from Basilea Pharmaceutica, MotifBio, Genentech, and Lysovant. D. M. is an employee of Nabriva Therapeutics and holds stock in Nabriva Therapeutics. S. N. is an employee of Johnson & Johnson and holds stock in Johnson & Johnson. J. H.-A. report grants from NIH (funding to ARLG, payments made to institution). S. B. D. reports grants or contracts from NIH funding to the ARLG; research funding from Gilead, Regeneron, and NIH; payments for clinical event adjudication committees from Basilea and Shionogi (pending); and consulting fees from Genentech. All remaining authors: No reported conflicts of interest.

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

- Boucher HW, File TM, Fowler VG, Jezek A, Rex JH, Outterson K. Antibiotic development incentives that reflect societal value of antibiotics. *Clin Infect Dis* **2021**; 72:e420–1.
- Árdal C, Balasegaram M, Laxminarayan R, et al. Antibiotic development—economic, regulatory and societal challenges. *Nat Rev Microbiol* **2020**; 18:267–74.
- 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.
- 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.
- Doernberg SB, Tran TTT, Tong SYC, et al. 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. *Clin Infect Dis* **2019**; 68:1691–8.
- Talbot GH, Powers JH, Hoffmann SC, et al. Developing outcomes assessments as endpoints for registrational clinical trials of antibacterial drugs: 2015 update from the Biomarkers Consortium of the Foundation for the National Institutes of Health. *Clin Infect Dis* **2016**; 62:603–7.
- Talbot GH, Das A, Cush S, et al. Evidence-based study design for hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. *J Infect Dis* **2019**; 219:1536–44.
- Knirsch C, Alemayehu D, Botgros R, et al. Improving conduct and feasibility of clinical trials to evaluate antibacterial drugs to treat hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia: recommendations of the Clinical Trials Transformation Initiative Antibacterial Drug Development Project Team. *Clin Infect Dis* **2016**; 63(Suppl 2):S29–36.
- van Duin D, Lok JJ, Earley M, et al. Colistin versus ceftazidime-avibactam in the treatment of infections due to carbapenem-resistant Enterobacteriaceae. *Clin Infect Dis* **2018**; 66:163–71.
- Lodise TP, Rosenkranz SL, Finnemeyer M, et al. The emperor's new clothes: prospective observational evaluation of the association between initial vancomycin exposure and failure rates among adult hospitalized patients with methicillin-resistant *Staphylococcus aureus* bloodstream infections (PROVIDE). *Clin Infect Dis* **2020**; 70:1536–45.
- van Duin D, Arias CA, Komarow L, et al. Molecular and clinical epidemiology of carbapenem-resistant Enterobacteriales in the USA (CRACKLE-2): a prospective cohort study. *Lancet Infect Dis* **2020**; 20:731–41.
- Williams DJ, Creech CB, Walter EB, et al. Short versus standard-course outpatient antibiotic therapy for community-acquired pneumonia in children: the SCOUT-CAP randomized clinical trial. *JAMA Pediatr* **2022**; 176:253–61.
- Kaye KS, Rice LB, Dane AL, et al. Fosfomycin for injection (ZTI-01) versus piperacillin-tazobactam for the treatment of complicated urinary tract infection including acute pyelonephritis: ZEUS, a phase 2/3 randomized trial. *Clin Infect Dis* **2019**; 69:2045–56.
- Portsmouth S, van Veenhuizen D, Echols R, et al. Cefiderocol versus imipenem-cilastatin for the treatment of complicated urinary tract infections caused by gram-negative uropathogens: a phase 2, randomised, double-blind, non-inferiority trial. *Lancet Infect Dis* **2018**; 18:1319–28.
- Naber KG, Llorens L, Kaniga K, Kotey P, Hedrich D, Redman R. Intravenous doripenem at 500 milligrams versus levofloxacin at 250 milligrams, with an option to switch to oral therapy, for treatment of complicated lower urinary tract infection and pyelonephritis. *Antimicrob Agents Chemother* **2009**; 53:3782–92.
- Chambers HF, Evans SR, Patel R, et al. Antibacterial Resistance Leadership Group 2.0: back to business. *Clin Infect Dis* **2021**; 73:730–9.
- US Department of Health and Human Services. Complicated urinary tract infections: developing drugs for treatment guidance for industry. 2018: 19. Available at: <https://www.fda.gov/files/drugs/published/Complicated-Urinary-Tract-Infections—Developing-Drugs-for-Treatment.pdf>. Accessed 21 December 2021.
- US Department of Health and Human Services. E6(R2)-good-clinical-practice-integrated-addendum-to-ICH-E6(R1).pdf. 2018. Available at: <https://www.fda.gov/files/drugs/published/E6%28R2%29-Good-Clinical-Practice-Integrated-Addendum-to-ICH-E6%28R1%29.pdf>. Accessed 21 December 2021.
- Halperin M, Hamdy MI, Thall PF. Distribution-free confidence intervals for a parameter of Wilcoxon-Mann-Whitney type for ordered categories and progressive censoring. *Biometrics* **1989**; 45:509–21.
- King HA, Doernberg SB, Miller J, et al. Patients' experiences with *Staphylococcus aureus* and gram-negative bacterial bloodstream infections: a qualitative descriptive study and concept elicitation phase to inform measurement of patient-reported quality of life. *Clin Infect Dis* **2021**; 73:237–47.