

UCSF

UC San Francisco Previously Published Works

Title

Novel Method Suggests Global Superiority of Short-Duration Antibiotics for Intra-abdominal Infections.

Permalink

<https://escholarship.org/uc/item/2z24s7h3>

Journal

Clinical Infectious Diseases, 65(9)

ISSN

1058-4838

Authors

Celestin, Arthur R
Odom, Stephen R
Angelidou, Konstantia
et al.

Publication Date

2017-10-16

DOI

10.1093/cid/cix569

Peer reviewed

Novel Method Suggests Global Superiority of Short-Duration Antibiotics for Intra-abdominal Infections

Arthur R. Celestin,¹ Stephen R. Odom,¹ Konstantia Angelidou,² Scott R. Evans,² Raul Coimbra,³ Christopher A. Guidry,⁴ Joseph Cuschieri,⁵ Kaysie L. Banton,⁶ Patrick J. O'Neill,⁷ Reza Askari,⁸ Nicholas Namias,⁹ Therèse M. Duane,¹⁰ Jeffrey A. Claridge,¹¹ E. Patchen Dellinger,⁵ Robert A. Sawyer,¹² and Charles H. Cook¹

¹Department of Surgery, Beth Israel Deaconess Medical Center, and ²Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts; ³Department of Surgery, University of California, San Diego Medical Center; ⁴Department of Surgery, University of Texas Medical Branch, Galveston; ⁵Department of Surgery, University of Washington School of Medicine, Seattle; ⁶Department of Surgery, University of Minnesota, Minneapolis; ⁷Department of Surgery, University of Arizona School of Medicine, Phoenix; ⁸Department of Surgery, Brigham and Women's Hospital, Boston, Massachusetts; ⁹Department of Surgery, University of Miami Health System, Florida; ¹⁰Department of Surgery, John Peter Smith Hospital, Fort Worth, Texas; ¹¹Department of Surgery, MetroHealth System, Case Western Reserve University, Cleveland, Ohio; and ¹²Department of Surgery, University of Virginia Health System, Charlottesville

Desirability of outcome ranking and response adjusted for duration of antibiotic risk (DOOR/RADAR) are novel and innovative methods of evaluating data in antibiotic trials. We analyzed data from a noninferiority trial of short-course antimicrobial therapy for intra-abdominal infection (STOP-IT), and results suggest global superiority of short-duration therapy for intra-abdominal infections.

Keywords. antibiotic resistance; desirability of outcome ranking; duration of antibiotic therapy; surgical infection; intra-abdominal infection.

(See the Editorial Commentary by Solomkin on pages 1580–1.)

The desire to optimize antibiotic duration has been well documented [1, 2]. Motivations for less antibiotic use include microbial resistance and cost, preventing unnecessary toxicity, and improving quality of life [3–5]. The Trial of Short-Course Antimicrobial Therapy for Intraabdominal Infection (STOP-IT) evaluated short-duration antibiotic therapy for treatment of abdominal infections after initial source control [6]. This multisite, randomized controlled trial of 518 patients concluded

that short, fixed-duration antibiotic therapy is noninferior to traditional, longer-duration therapy until resolution of physiologic abnormalities.

Designing clinical trials to effectively evaluate antibiotic therapy is challenging due to inherent biases of noninferiority trials that require very large sample sizes. Like many such trials, STOP-IT had difficulty reaching enrollment targets. Part of the reason that large sample sizes are required is that the endpoints are binary (eg, “success” or “failure”). Moreover, more complex definitions of outcomes that incorporate recovery, adverse events, secondary infections, and the risk for antibiotic resistance are lacking. As a practical consequence, clinicians may be reluctant to try new strategies tested only by noninferiority trials [7].

Recently, desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk analyses (RADAR) have been introduced as innovative approaches [8]. The concept of DOOR is to rank all trial participants with ordinal clinical outcomes that include benefits, harms, and quality of life. For analysis, patients in control and experimental groups are categorized and DOOR distributions are compared. Probabilities that a patient will have a better DOOR score if assigned to a treatment strategy are calculated. For this calculation, a probability of 50% means that the odds of an experimental group having a better outcome than a control patient are equivalent to a coin toss. Likewise, because probabilities from this analysis have a sum of 1, if an experimental treatment has a DOOR probability of 70%, then the control treatment will have a probability of 30%. In that case, patients assigned to experimental treatment are 70% more likely to have a better outcome than those assigned to control treatment. These analyses may be viewed as systematic analyses based on the totality of clinical outcomes.

RADAR further ranks patients with similar clinical outcomes by durations of antibiotic therapy. RADAR is thus used to “break ties” between patients with the same clinical outcome ranking. For example, a patient with clinical benefit and no adverse effects who is on antibiotics for 3 days is ranked higher than a patient with the same clinical outcome but who is on antibiotics for 6 days. Probabilities of more desirable outcomes are then calculated based on these adjusted ranks. The goal is to avoid complexities associated with noninferiority trials while simultaneously potentially reducing required sample sizes.

We retrospectively applied this method to STOP-IT data to estimate probabilities of desirable outcomes for short-course antibiotic therapy for intra-abdominal infections. Our hypothesis was that DOOR/RADAR would show superiority of short-course antibiotic therapy vs traditional therapy for intra-abdominal infection. We further sought to determine if such conclusions might be reached with even smaller sample sizes.

Received 6 January 2017; editorial decision 3 April 2017; accepted 20 June 2017; published online August 28, 2017.

Presented in part: 36th Annual Meeting of the Surgical Infection Society, Palm Beach, Florida, 19–21 May 2016.

Correspondence: A. R. Celestin, Department of Surgery, Beth Israel Deaconess Medical Center, 110 Francis St, Ste 3A, Boston, MA 02215-5501 (acelesti@bidmc.harvard.edu).

Clinical Infectious Diseases® 2017;65(9):1577–9

© The Author 2017. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/cid/cix569

METHODS

We applied DOOR analysis to results from all 518 STOP-IT trial patients. Patients were categorized as having the following mutually exclusive outcomes: (1) recovery, with no complications; (2) recovery, with extraabdominal infection (including *Clostridium difficile*); (3) recovery with surgical site infection/wound infection; (4) recovery with recurrent intra-abdominal infection requiring procedure; or (5) death. Treatment failures requiring direct intervention were considered worse than skin and soft tissue infections. For all patients in the STOP-IT trial, the probability that a randomly selected patient would have a better DOOR if assigned to short-course (SC) vs traditional duration (TD) was estimated using confidence intervals (CIs). Next, RADAR was applied to DOOR results to further stratify patients with similar outcomes. To compare SC vs TD based only on clinical outcomes (preventing undue influence of antibiotic duration), DOOR probabilities for experimental and control groups were additionally calculated without including antibiotic duration (without RADAR).

RESULTS

For all-comers, 71.98% (n = 185) of SC patients and 73.08% (n = 190) of TD patients were in category 1; 6.61% (n = 17) of SC and 4.62% (n = 12) of TD were category 2; 4.67% (n = 12) of SC and 8.08% (n = 21) of TD were category 3; 15.56% (n = 40) of SC and 13.46% (n = 35) of TD were category 4; and 1.17% (n = 3) of SC and 0.77% (n = 2) of TD were category 5. Analyses based on this 5-tiered scheme showed that the probability that a randomly selected patient would have a better DOOR score if receiving SC antibiotics was 49.33% (95% CI, 46.20%–54.44%). Conversely, the probability that a random patient undergoing TD therapy would have a better DOOR score was 50.67%.

When RADAR was applied, the probability that randomly selected patients would have a better DOOR score if receiving SC antibiotics was 63.64% (95% CI, 58.63%–68.69%).

We next determined if similar conclusions might be possible with even smaller sample sizes. For this we analyzed the first 150 patients enrolled in STOP-IT (75 per group) using DOOR/RADAR. For this smaller group, the probability of an improved DOOR was 66.3% for patients in the SC group.

DISCUSSION

In this retrospective DOOR/RADAR analysis, patients receiving short-course therapy had higher probabilities of more desirable outcomes than those receiving a traditional duration of therapy. This extends and strengthens the original conclusions of STOP-IT. Whereas STOP-IT was a multicenter, randomized controlled study that was projected to require a sample size of 1010 patients to show noninferiority based on a single endpoint, our results suggest global superiority of short-course antibiotics even with the sample size (518) that was accrued.

Finally, our results show that DOOR/RADAR analysis would have suggested superiority of short-course antibiotics after enrollment of only 150 patients.

A critical component to these analyses is meaningful ranking schemes, and creating appropriate clinical outcome rankings has the potential to be challenging (see Figure 1). It is important that researchers carefully choose predetermined, consensus-driven, ordinal levels, as this is probably the most important factor in ensuring the applicability/validity of this method. It is our hope that the proposed ranking scheme may be useful for future studies of patients undergoing treatment for surgical infections. Further analysis and broader consensus may improve this ranking scheme and allow more meaningful outcome analyses, with similar studies of other existing datasets serving as important next steps.

If the goal is to quantify the global experience of patients, then ultimately patients themselves should probably contribute to desirability rankings. Despite our clinical opinions, it is possible that a patient with *Clostridium difficile* might be far more uncomfortable and require more treatment than a patient with intraabdominal abscess. Moreover, desirability ranking analyses might be even more useful in other applications, such as cancer therapy, where duration of treatment and complications can be included with current binary survival analyses typically used in such trials.

There are limitations inherent to the application of DOOR/RADAR to this trial. First, the global negative effects of longer antibiotic use have not been quantified, and more studies that quantify benefits and harms of antibiotic duration are clearly needed. Because of its retrospective nature, clinical outcomes for the current study could only be chosen based on data points collected for STOP-IT. This runs the risk of skewing/heavily weighing outcome scores toward the beliefs driving the variable(s) picked for RADAR. Prospectively constructed trials will hopefully help avoid this bias. In addition, rank-based analyses may be unable to elucidate a less-common but more important clinical deficit due to a more frequent but less important advantage in antibiotic use. For DOOR/RADAR to stand alone, thorough analyses of component outcomes would be required. Trials originally powered for DOOR/RADAR superiority would require larger numbers to evaluate component outcomes, obviating the sample size advantage. Alternatively, investigators can consider separate analyses on component outcomes that comprise the DOOR, and even partial credit analyses (discussed in detail in [9]) to avoid skewing outcomes. Given the possible statistical pitfalls, it is crucial that this method always be accompanied by additional analyses, including those that ignore RADAR.

In summary, DOOR/RADAR analysis can be used in antibiotic trials to globally evaluate for superiority of new antibiotic strategies that previously might have been reportable only as noninferior. The current report supports the conclusions of

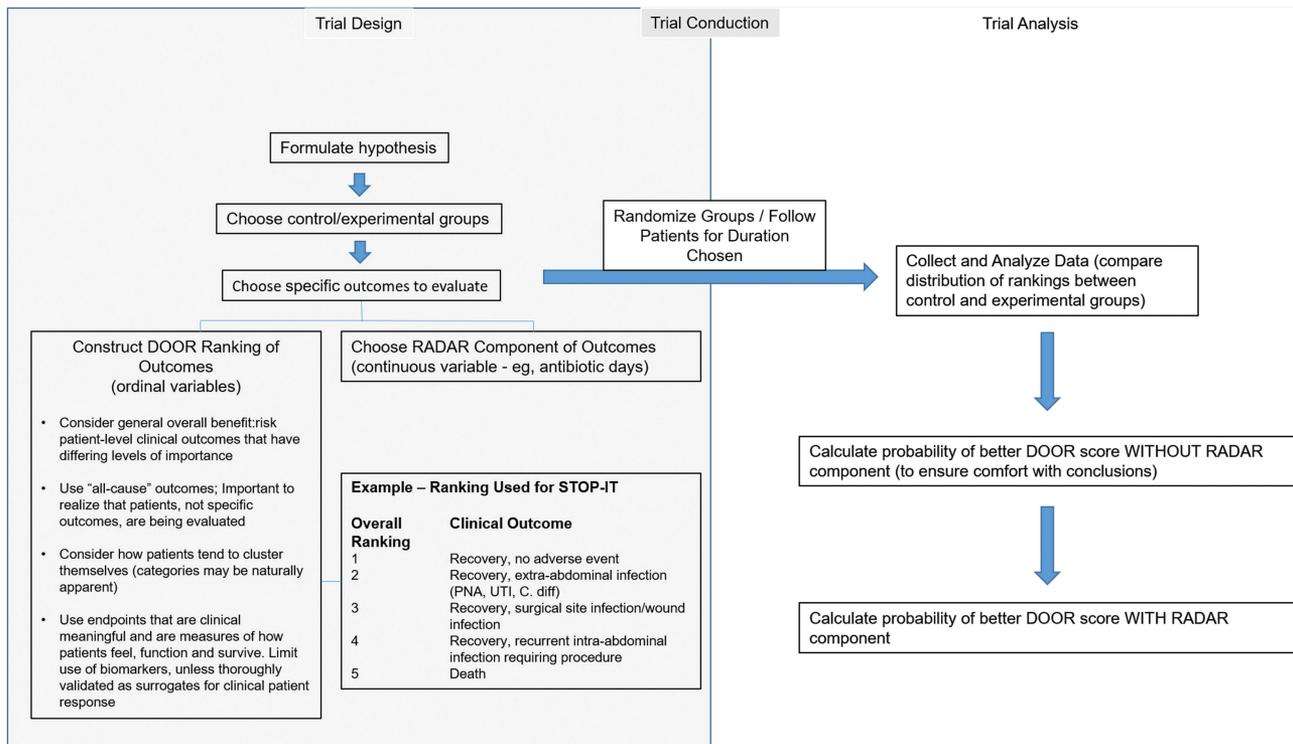


Figure 1. Overview of suggested trial design and analysis in clinical trials using DOOR/RADAR. Abbreviations: DOOR/RADAR, desirability of outcome ranking/response adjusted for duration of antibiotic risk; PNA, pneumonia; STOP-IT, Trial of Short-Course Antimicrobial Therapy for Intraabdominal Infection; UTI, urinary tract infection.

STOP-IT, and suggests further that a short duration of antibiotic therapy is superior to a longer duration of therapy for complicated intra-abdominal infection. We also propose a novel surgical infection outcomes classification system for consideration in future studies. By careful prospective choice of ordinal ranking levels and the use of the supporting analyses described above, we believe that future studies will confirm DOOR/RADAR as an advantageous methodology that allows stronger conclusions to be drawn from smaller sample sizes.

Notes

Financial support. R. A. S., R. C., T. M. D., J. A. C., and P. J. O. received funding from the National Institutes of Health (NIH).

Potential conflicts of interest. S. R. E. has received personal fees from Takeda/Millennium; Pfizer; Roche; Novartis; Achaogen; Huntington's Study Group; Auspex; Alcon; Merck; Chelsea; Mannkind; QRx Pharma; IMMPECT; Genentech; Affymax; FzioMed; Amgen; GSK; Sunovion; 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; PPRECISE; Statistical Communications in Infectious Diseases; Cubist; AstraZeneca; Teva; Repros; Austrian Breast and Colorectal Cancer Study Group/Breast International Group and the Alliance Foundation Trials; Zeiss; Dexcom; American Society for Microbiology; and Taylor and Francis. R. A. S. has received consulting fees from 3M; Merck; Pfizer; and GlaxoSmithKline. T. M. D. has received payment for providing expert testimony for private law firms. E. P.

D. has received payment for consulting from 3M; Therevance; and Melinta; for providing expert testimony for various private malpractice attorneys; and for giving lectures at Fraser Health (Vancouver, B.C.); Washington University; the University of California; Irvine; and the Washington State Hospital Association. P. J. O. has received consulting fees from the data and safety monitoring board of ACI Clinical (Bala Cynwyd, Pennsylvania). 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

- Luyt CE, Bréchet N, Trouillet JL, Chastre J. Antibiotic stewardship in the intensive care unit. *Crit Care* **2014**; *18*:480.
- 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.
- Rice LB. The clinical consequences of antimicrobial resistance. *Curr Opin Microbiol* **2009**; *12*:476–81.
- Gould IM. Coping with antibiotic resistance: the impending crisis. *Int J Antimicrob Agents* **2010**; *36*(suppl 3):S1–2.
- Bignardi GE. Risk factors for *Clostridium difficile* infection. *J Hosp Infect* **1998**; *40*:1–15.
- Sawyer RG, Claridge JA, Nathens AB, et al. Trial of short-course antimicrobial therapy for intraabdominal infection. *N Engl J Med* **2015**; *372*:1996–2005.
- Fleming TR. Current issues in non-inferiority trials. *Stat Med* **2008**; *27*:317–32.
- 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.
- 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.