Milken Institute School of Public Health
THE GEORGE WASHINGTON UNIVERSITY

THE BIOSTATISTICS CENTER

ARLG
Antibacterial Resistance Leadership Group

THE DOOR IS OPEN

Web tools for patient-centric, pragmatic benefit:risk evaluation in clinical trials

Toshimitsu Hamasaki, PhD, MS, Pstat®

The Biostatistics Center | Department of Biostatistics and Bioinformatics
Milken Institute School of Public Heath | The George Washington University

On be half of the GWU DOOR Research Team Weixiao Dai; Guoqing Diao; Scott R. Evans; Yijie He, Lizhao Ge; Richard Shu; Qihang Wu; Shanshan Zhan

The DOOR is open

Outline

- The Desirability of Outcome Ranking Methodology (DOOR) methodology: Motivation
- Development of DOOR outcomes
- Online tool for DOOR analyses
- Online tool for clinical trial designs
- Summary

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number UM1AI104681. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The Desirability Of Outcome Ranking (DOOR) methodology

Patient-centric, benefit:risk evaluation

- A paradigm for the design, monitoring, analysis, interpretation and reporting of clinical trials and other research studies based on patient-centric benefit:risk evaluation (Evans et al 2015; Evans and Follmann 2016).
  - "**Using Outcomes to Analyze Patients**" rather than "**Patients to Analyze Outcomes**"
  - Motivated by how to answer the most important question in treating patients in clinical practice

TRT A

Frequency		Efficacy: Yes	Efficacy: No
Toxicity	Yes	50	0
	No	0	50

50/100 =50%

50/100=50%

TRT B

		Efficacy: Yes	Efficacy: No
Toxicity	Yes	25	25
	No	25	25

50/100 =50%

50/100=50%

TRT C

		Efficacy: Yes	Efficacy: No
Toxicity	Yes	0	50
	No	50	0

50/100 =50%

50/100=50%

Evans SR et al. Clin Infect Dis 2015; 61:699-806. Evans SR, Follmann D. Stat Biopharm Res 2016; 8:386-393

The DOOR outcomes

Development of DOOR outcomes

- **Overall ordinal composite outcome of important clinical outcomes**
  - Tradeoffs among outcomes
  - Cumulative nature of benefits and harms on patients
- **The ARLG Innovation Working Group**
  - Proposed the DOOR outcomes for ABSSSI; Bacteremia; cIAI; cUTI; HABP/VABP
  - Applied DOOR outcomes to registrational trials in cUTI and HABP/VAB (Howard-Anderson et al. 2023a, b)
  - Collaborated with FDA Antibacterial Drug Resistance (DOOR) Fellowship: Evaluated the DOOR based on data from registrational trials in cIAI submitted to FDA (Kinamon et al. 2023)

- Absence of Clinical Response
- Infectious Complications
- Serious Adverse Events (SAEs)
- Death

Howard-Anderson J et al. Clin Infect Dis 2023a; 76:1157-1165. Howard-Anderson J et al. Clin Infect Dis 2023b: ciad5760. Kinamon T et al. Clin Infect Dis 2023;77:649-656

DOOR outcome analyses

ARLG recommendations: Simple, robust approach

Analysis	Outcome	Statistical method
Descriptive analysis	• DOOR • Components	• Summary distribution table by intervention group • Bar-chart by intervention group
	• DOOR and Components	• Anthology of Patient Stories (APS) plot
Rank-based analysis: DOOR probability	• DOOR • Components	• Forest Plot of estimates of the DOOR probability for the DOOR and respective components
	• DOOR	• Forest plot of the estimates for the cumulative DOOR probability based on sequential dichotomization of the DOOR outcome
Grade-based Analysis: Partial Credit	• DOOR	• Welch's t-statistic based analysis • Scatter plot of the differences in mean partial credit between interventions against the corresponding DOOR probabilities

Online tools for implementing DOOR analyses

DOOR apps

	Standard Edition	Professional Edition
Data Input	Summary table by group	Individual patient-level data
Analysis		
1. Descriptive analysis Summary table Bar-chart Anthology of patient stories plot	 ✅ ✅	 ✅ ✅ ✅
2. Rank-based analysis DOOR prob forest plot Dichotomized DOOR prob forest plot	 ✅ ✅	 ✅ ✅
3. Grade-based analysis Partial credit analysis summary Partial credit vs DOOR prob plot Partial credit forest plot	 ✅ ✅ ✅	 ✅ ✅ ✅
4. Tie-breaker analysis		✅
5. Inverse probability weighting		✅
Labels customization, Data save	✅	✅

Online tools for implementing DOOR analyses: Standard edition

Autofill the ARLG-proposed or other DOOR outcomes

DOOR Analyses: Standard Edition | Data Input Table | DOOR Distribution Summary Tabl

Pre-specified Settings

Default

Data Format

Frequencies (N) | Percentages (%)

# of DOOR Ranks (Maximum: 10)

5

# of DOOR Components (Maximum: 10)

4

Test Intervention Label

Treatment

Control Intervention Label

Control

Method for Confidence Interval (CI)

Halperin et al (1989)

Pseudo-Score Approach for Halperin et al (1989)

Confidence Level for Two-sided Confidence Interval

0.5 0.55 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 0.99

0.95

Unit for Expected Gained (+) or Loss (-)

1000

# of Grading keys (Maximum: 7)

1

Labels for Grade Keys

DOOR (Most desirable to

(%)

0)

Pre-specified Settings

Default

Default

ARLG

cUTI; HABP/VABP; ABSSSI; Bacteremia

Prioritized efficacy; Prioritized safety

Phage

cIAI (FDA)

HABP/VABP (FDA)

STROKE

Modified Rankin Scale for Neurologic Disability (6-level)

Modified Rankin Scale for Neurologic Disability (7-level)

CANCER

Karnofsky Performance Status Scale

EXAMPLES

DORI-05

ACTT

Online tools for implementing DOOR analyses: Standard edition

DOOR apps: Data Input

DOOR Analyses: Standard Edition | Data Input Table | DOOR Distribution Summary Table | DOOR Forest Plot | Partial Credit Analysis | Support | Logs

Pre-specified Settings

Default

Data Format

Frequencies (N) Percentages (%)

# of DOOR Ranks (Maximum: 10)

5

# of DOOR Components (Maximum: 10)

4

Test Intervention Label

Treatment

Control Intervention Label

Control

Method for Confidence Interval (CI)

Halperin et al (1989)

Pseudo-Score Approach for Halperin et al (1989)

Confidence Level for Two-sided Confidence Interval

0.5 0.95

0.5 0.55 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 0.99

Unit for Expected Gained (+) or Loss (-)

1000

# of Grading keys (Maximum: 7)

1

Labels for Grade Keys

DOOR Distribution by Intervention

DOOR (Most desirable to least desirable) Rank	Doripenem	Levofloxacin
Alive with no events	263	253
Alive with 1 event	93	111
Alive with 2 events	16	9
Alive with 3 events	1	1
Death	1	0
Total (N)	374	374

DOOR Components Distribution by Intervention

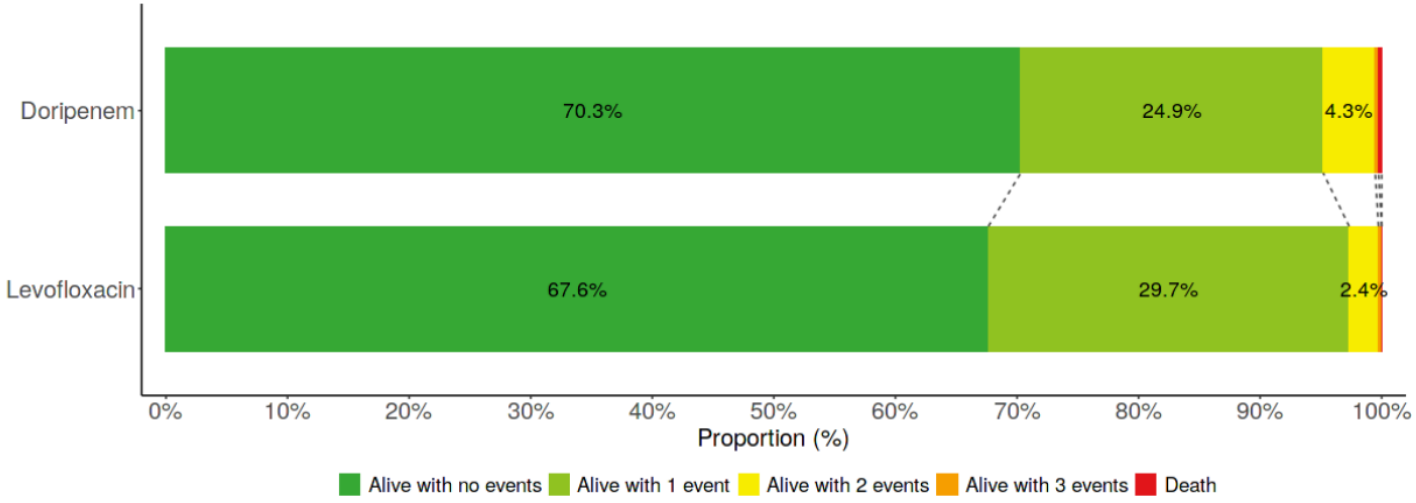
DOOR Component	Doripenem	Levofloxacin
Clinical Failure	81	113
Infectious Complications	23	5
SAEs	25	14
Death	1	0

Howard-Anderson J, et al. Improving Traditional Registrational Trial End Points: Development and Application of a Desirability of Outcome Ranking End Point for Complicated Urinary Tract Infection Clinical Trials. Clin Infect Dis 2023 76:1157-1165.

An Illustration: DORI05- Doripenem vs Levofloxcin in cUTI

Descriptive analysis: DOOR outcome distribution by intervention group

	Doripenem				Levofloxacin					
			Cumulative				Cumulative		Expected Gained (+) or Loss (-) (per1000)	
DOOR	n	(%)	n	(%)	n	(%)	n	(%)	Per Category	Cumulative
Alive with no events	263	70.3	263	70.3	253	67.6	253	67.6	27	27
Alive with 1 event	93	24.9	356	95.2	111	29.7	364	97.3	-48	-21
Alive with 2 events	16	4.3	372	99.5	9	2.4	373	99.7	19	-3
Alive with 3 events	1	0.3	373	99.7	1	0.3	374	100.0	0	-3
Death	1	0.3	374	100.0	0	0.0	374	100.0	3	0
Total (N)	**374**				**374**					

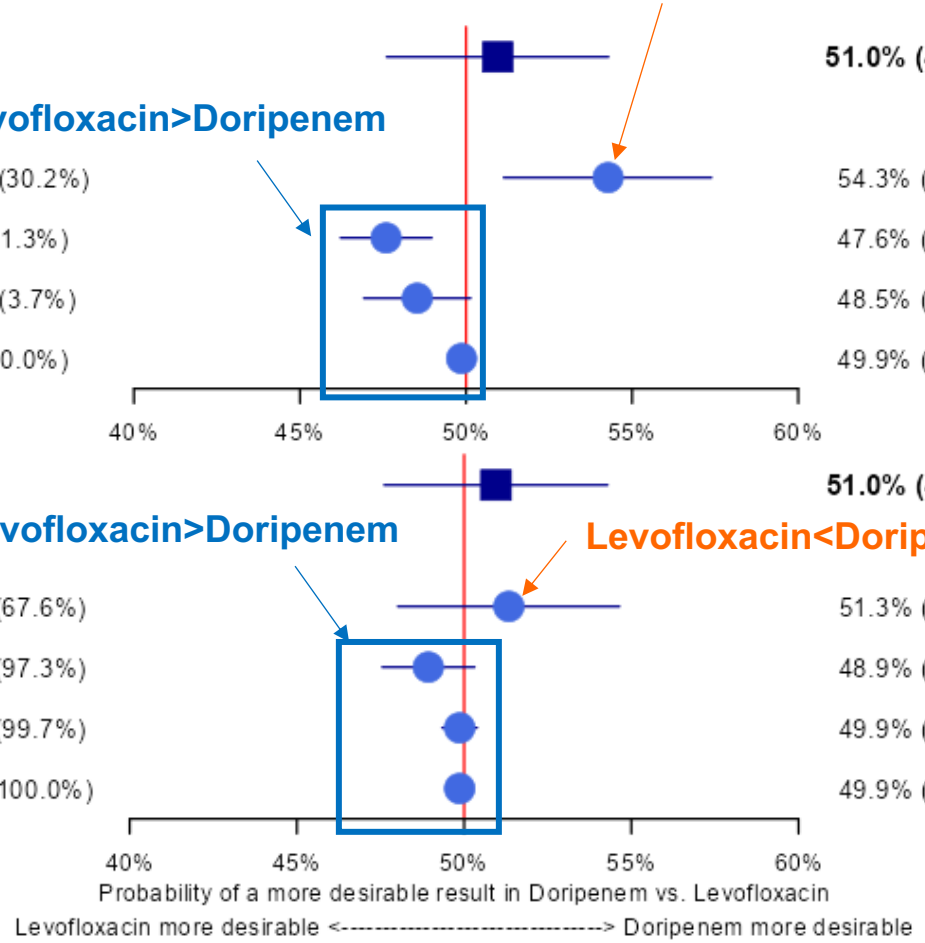


Howard-Anderson J, et al. Improving Traditional Registrational Trial End Points: Development and Application of a Desirability of Outcome Ranking End Point for Complicated Urinary Tract Infection Clinical Trials. Clin Infect Dis 2023 76:1157-1165.

An Illustration: DORI05

Rank-based analysis: Forest plot of the DOOR and respective components

	Doripenem (N=374) n(%)	Levofloxacin (N=374) n(%)	DOOR probability (95% CI)	p-value
DOOR			**51.0% (47.6%, 54.3%)**	**0.5758**
DOOR Component				
Clinical Failure	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%)	
SAEs	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%)	
DOOR			**51.0% (47.6%, 54.3%)**	**0.5758**
Sequential Dichotomization of DOOR				
DOOR <= 1 - DOOR > 1	263 (70.3%)	253 (67.6%)	51.3% (48.0%, 54.6%)	
DOOR <= 2 - DOOR > 2	356 (95.2%)	364 (97.3%)	48.9% (47.5%, 50.3%)	
DOOR <= 3 - DOOR > 3	372 (99.5%)	373 (99.7%)	49.9% (49.3%, 50.4%)	
DOOR <= 4 - DOOR > 4	373 (99.7%)	374 (100.0%)	49.9% (49.5%, 50.2%)	



Howard-Anderson J, et al. Improving Traditional Registrational Trial End Points: Development and Application of a Desirability of Outcome Ranking End Point for Complicated Urinary Tract Infection Clinical Trials. Clin Infect Dis 2023 76:1157-1165.

An Illustration: DORI05

Grade-based analysis: Partial credit analysis summary

DOOR (Most desirable to least desirable)	Grading key 1	Grading key 2	Grading key 3	Grading key 4	Grading key 5
Alive with no events	100	100	100	100	100
Alive with 1 event	100	100	100	0	80
Alive with 2 events	100	100	0	0	60
Alive with 3 events	100	0	0	0	40
Death	0	0	0	0	0

Statistics	DOR	LEV	DOR	LEV	DOR	LEV	DOR	LEV	DOR	LEV
Mean (SD)	99.7(5.2)	100.0(0.0)	99.5 (7.3)	99.7 (5.2)	95.2 (21.4)	97.3 (16.2)	70.3 (45.7)	67.6 (46.8)	92.9 (12.4)	92.9 (10.8)
Diff. in means(95%CI)	-0.3 (-0.8, 0.3)		-0.2 (-1.2, 0.6)		-2.1 (-4.9 , 0.6)		2.7 (-4.0 , 9.3)		0.0 (-1.7 , 1.6)	
P-value	0.3180		0.5635		0.1237		0.4299		0.9500	
DOOR probability (%) (95%CI)	49.9 (49.5 , 50.2)		49.9 (49.3, 50.4)		48.9 (47.5, 50.3)		51.3 (48.0, 54.6)		51.0 (47.6, 54.3)	
P-value	0.3173		0.5632		0.1236		0.4296		0.5758	

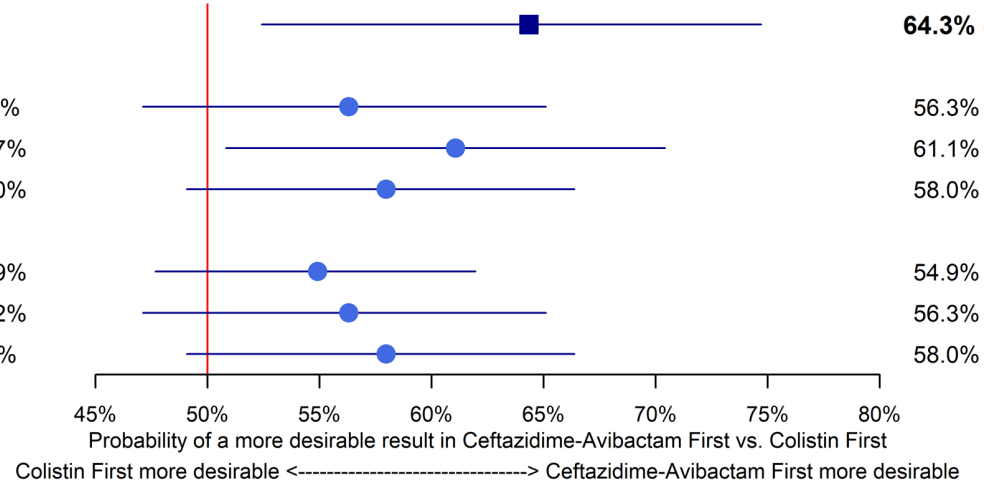
DOR: Doripenem; LEV: Levofloxacin

An Illustration: CRACKE I- Colistin versus Ceftazidime-Avibactam in CRE

IPW analysis using the Professional Edition

Inverse Probability Weighted (IPW) Analysis

	Ceftazidime-Avibactam First (N=26) %	Colistin First (N=46) %	DOOR probability (95% CI)	p-value
DOOR			**64.3% (52.4%, 74.7%)**	**0.0229**
Sequential Dichotomization of DOOR				
DOOR <= 1 - DOOR > 1	20.5%	7.8%	56.3% (47.1%, 65.1%)	
DOOR <= 2 - DOOR > 2	85.9%	63.7%	61.1% (50.8%, 70.4%)	
DOOR <= 3 - DOOR > 3	90.9%	75.0%	58.0% (49.1%, 66.4%)	
DOOR Components				
Renal failure	5.1%	14.9%	54.9% (47.7%, 62.0%)	
Not discharged	79.5%	92.2%	56.3% (47.1%, 65.1%)	
Death	9.1%	25%	58.0% (49.1%, 66.4%)	



van Duin D et al. Colistin versus Ceftazidime-Avibactam in the Treatment of Infections Due to Carbapenem-Resistant Enterobacteriaceae. Clin Infect Dis. 2018; 66:163-171. **Rank 1**. Discharged home; **Rank 2**. Alive in hospital or discharged not to home, no incident renal failure; **Rank 3**. Alive in hospital or discharged not to home, incident renal failure: **Rank 4**. Hospital death

Designing a clinical trial with DOOR methodology

A tool for power and sample size assessment: Data input

DOOR: Power and Sample Size Assessment ≡

Assessment

DOOR Probability to Be Detected

Configurations/Settings

One or Two-sided Test

- One-sided
- Two-sided

Significance Level (α)

0.05

Allocation Ratio

0.5

Desired Power (1-β) (%)

80

DOOR Probability of Null Hypothesis (%)

50

Method

- Iterative Method
- Method by Tang (2011)
- Method by Noether (1987)

Total Sample Size | **Test** | **Control**

Power Evaluation by Simulation

- No
- Yes

ted

ined by

ategory Proportions

ervention

t desirable)

Calculated DOOR Probability (%)

59.0

ontrol

0

0

0

00

Designing a clinical trial with DOOR methodology

A tool for power and sample size assessment: Output

Outputs

Compute | Delete | Save

Empirical Power Given Sample Size	# of Dataset Generations	Seed for Dataset Generations	Report
NA%	NA	NA	Generate

Download

Sample Size / Power Report

A sample size of **131** in **Test** and **131** in **Control** (in total **262**) has **80.3**% power to reject the null hypothesis: the DOOR probability = **50.0**%, assuming that a value of DOOR probability to be detected is **59.0**%, based on the proportions of DOOR outcome with **3** ranks shown below, for a **two-sided** Wilcoxon-Mann-Whitney test at **5.0**% significance level, using the method in **Tang (2011)**.

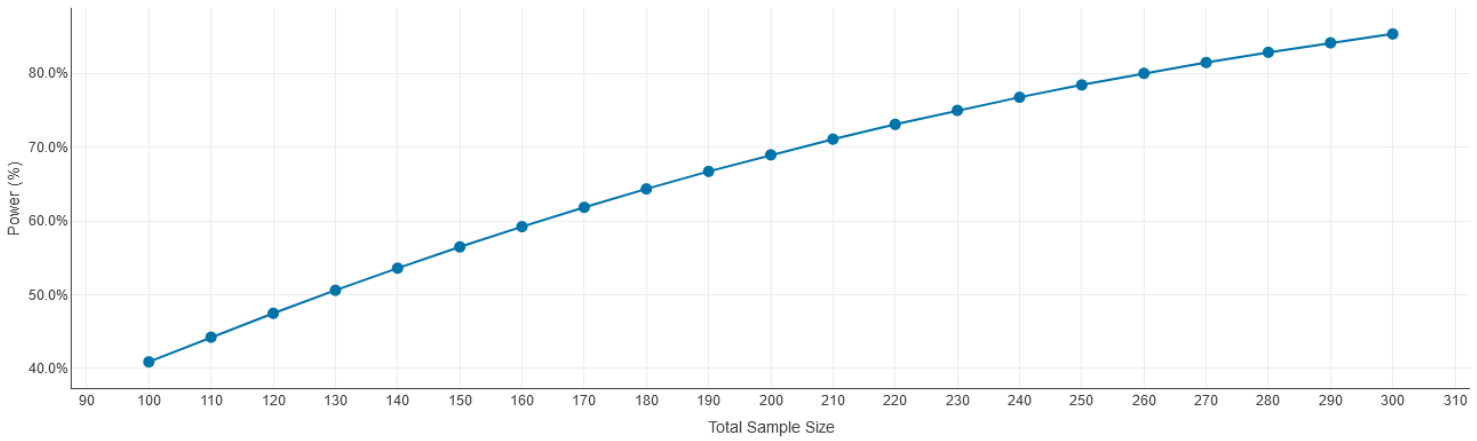
Power vs. Total Sample Size | Power vs. DOOR Probability

Range of Total Sample Size: 100 to 300

Increment by: 10

Create

Download



Total Sample Size	Power (%)
100	40.9
110	44.2
120	47.5
130	50.6
140	53.6
150	56.5
160	59.2
170	61.8
180	64.3
190	66.7

Previous 1 2 3 Next

Summary

DOOR methodology

- **DOOR outcome**: A global composite benefit:risk outcome at individual patients level, constructed on the basis of important clinical outcomes
- **Analyses: Simple, Robust approach**
  - **Rank-based analysis approach:** DOOR probability Pairwise comparison at individual patient level
  - **Grade-based analysis approach:** Evaluation of the impact of interventions based on patients' personal perspectives on the desirability of the DOOR outcome categories
    - Visualizes the impact of each category on the DOOR outcomes
    - Can incorporate patient preferences into treatment selections

Summary

Online tools for DOOR Methodology

- Statistical methods for analyzing DOOR outcomes require mathematical sophistication and knowledge of programming techniques, which can be a barrier for non-statisticians.
- A series of interactive web-based tool provide comprehensive tool for clinical researchers to implement DOOR methodology for their studies
- **More to come!**
  - Monitoring of clinical trials, including group-sequential and adaptive designs
  - Integrated analyses: meta-analysis
  - Covariate-adjusted analysis: stratified analysis
  - Subgroup analysis
  - Longitudinal time-to-event type DOOR outcomes

Milken Institute School of Public Health
THE GEORGE WASHINGTON UNIVERSITY
THE BIOSTATISTICS CENTER

ARLG
Antibacterial Resistance Leadership Group

The DOOR is Open!

THANK YOU

Toshimitsu Hamasaki

@ToshimitsuHama1
thamasaki@gwu.edu
https://methods.bsc.gwu.edu/

BACKUP

DOOR analyses

Concerns on common analyses for ordinal outcomes in clinical trials

Analysis	Feature	Concern
Dichotomized analysis	● An ordinal outcome is dichotomized to a binary outcome with a specified threshold (Ex. responder vs non-responder) ● Logistic-regression is then used to estimate the odds ratio and associated confidence interval of responder between groups as a measure of the treatment effect.	● May be inefficient from the statistical perspective due to the loss of information from ignoring finer but important gradations of patient status. ● May lead to decreased power or a necessary sample size increase to maintain power
Regression model based analysis	● A proportional-odds regression model is used to estimate the odds ratio and associated confidence interval across all the categories (common odds ratio) as a measure of the treatment effect	● Fail to provide intuitive interpretations, which helpful for clinical decision-making. ● Require the model's assumptions, sometimes strong assumptions to hold in order for model-based inferences to be valid.

Creating a DOOR outcome

ARLG proposed DOOR outcomes

- Absence of Clinical Response
- Serious Adverse Events
- **Infections complications**
- Death

If an newly developed infectious complication is an SAE, then the event is counted twice in deriving the DOOR outcome.

Disease	Infectious Complications
ABSSSI	Unplanned surgical for progression/ complication of original infection; Bacteremia; Septic shock; Osteomyelitis; *c.diff*
Bacteremia	Septic shock; Prolonged bacteremia on Day 5; Supportive complications or monastic site(s) of infection; *c.diff*
cIAI	Bacteremia; Septic shock; Peritonitis; Unplanned surgical for progression/ complication of original infection; *c.diff*
cUTI	Renal or intra-abdominal abscess; Septic shock; Bacteremia; Unplanned surgical for progression/ complication of original infection; *c.diff*
HABP/VABP	Complicated pleural effusion; Lung abscess/necrotizing pneumonia; ARDS; Meningitis; Bacteremia; Septic shock; Need for intubation; *c.diff*