ARLG
Antibacterial Resistance Leadership Group

Pragmatic Benefit:Risk Evaluation:
Healthy Disruption for Clinical Trials and Diagnostic Studies

Scott Evans, MS, PhD
Harvard University

October, 2017

ARLG
Antibacterial Resistance Leadership Group

Most clinical trials fail to provide the evidence needed to inform medical decision-making. However, the serious implications of this deficit are largely absent from public discourse.

DeMets and Califf, *JAMA*, 2011

If I had one hour to solve a problem, I would spend 55 minutes defining and understanding the problem, and 5 minutes solving it.

Albert Einstein

Question 1

- Suppose we measure the duration of hospitalization
- Shorter duration is better … or is it?
- The faster the patient dies, the shorter the duration
- Interpretation of an outcome needs context of other clinical outcomes for the same patient

Question 2

- We define analysis populations
  - Efficacy: ITT population
  - Safety: safety population
- Efficacy population ≠ safety population
- We combine these analyses into benefit:risk analyses
- To whom does this analysis apply?

Question 3

- Suppose a loved one is diagnosed with a serious disease
- You are selecting treatment
- 3 treatment options: A, B, and C
- 2 outcomes, equally important
  - Treatment success: yes/no
  - Safety event: yes/no

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

Which treatment would you choose?

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

Which treatment would you choose?

They all have the same success rate.

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

Which treatment would you choose?

They all have the same success rate.

A has the lowest safety event rate.

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

Which treatment would you choose?

They all have the same success rate.

A has the lowest safety event rate.

B and C are indistinguishable.

RCT Comparing A, B, and C
Analysis of Outcomes

A (N=100)	B (N=100)	C (N=100)
Success: 50%	Success: 50%	Success: 50%
Safety event: 30%	Safety event: 50%	Safety event: 50%

Which treatment would you choose?

They all have the same success rate.

A has the lowest safety event rate.

B and C are indistinguishable.

Choose A…right?

Analysis of Patients: 4 Possible Outcomes

		A (N=100) Success: 50% Safety event: 30%		B (N=100) Success: 50% Safety event: 50%		C (N=100) Success: 50% Safety event: 50%	
		Success		Success		Success	
		+	-	+	-	+	-
SE	+	15	15	50	0	0	50
	-	35	35	0	50	50	0

Analysis of Patients: 4 Possible Outcomes

		A (N=100) Success: 50% Safety event: 30%		B (N=100) Success: 50% Safety event: 50%		C (N=100) Success: 50% Safety event: 50%	
		Success		Success		Success	
		+	-	+	-	+	-
SE	+	15	15	50	0	0	50
	-	35	35	0	50	50	0

Analysis of Patients: 4 Possible Outcomes

		A (N=100) Success: 50% Safety event: 30%		B (N=100) Success: 50% Safety event: 50%		C (N=100) Success: 50% Safety event: 50%	
		Success		Success		Success	
		+	-	+	-	+	-
SE	+	15	15	50	0	0	50
	-	35	35	0	50	50	0

Analysis of Patients: 4 Possible Outcomes

		A (N=100) Success: 50% Safety event: 30%		B (N=100) Success: 50% Safety event: 50%		C (N=100) Success: 50% Safety event: 50%	
		Success		Success		Success	
		+	-	+	-	+	-
SE	+	15	15	50	0	0	50
	-	35	35	0	50	50	0

ARLG
Antibacterial Resistance Leadership Group

Our culture is to use patients to analyze the outcomes.

ARLG
Antibacterial Resistance Leadership Group

Our culture is to use patients to analyze the outcomes.

Shouldn't we use outcomes to analyze the patients?

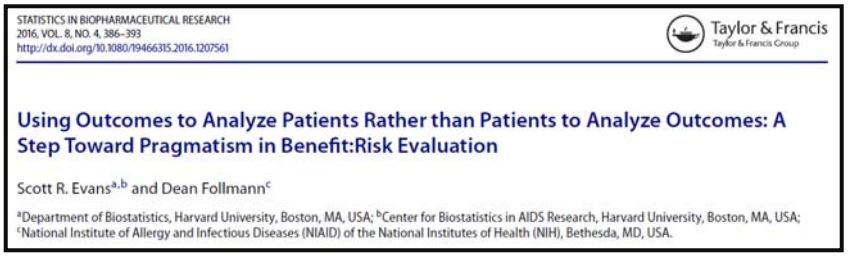
Scott's father (a math teacher) to his confused son many years ago:

"The order of operations is important…"

ARLG
Antibacterial Resistance Leadership Group

Instead of getting a great solution to the wrong problem, let's get a good solution to the right problem.

A Vision



STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2016, VOL. 8, NO. 4, 386–393
http://dx.doi.org/10.1080/19466315.2016.1207561

Taylor & Francis
Taylor & Francis Group

Using Outcomes to Analyze Patients Rather than Patients to Analyze Outcomes: A Step Toward Pragmatism in Benefit:Risk Evaluation

Scott R. Evans[a,b] and Dean Follmann[c]

[a]Department of Biostatistics, Harvard University, Boston, MA, USA; [b]Center for Biostatistics in AIDS Research, Harvard University, Boston, MA, USA; [c]National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH), Bethesda, MD, USA.

The good physician treats the disease.
The great physician treats the patient.

William Osler

Perhaps we should analyze the patient.

The Future of Clinical Trials

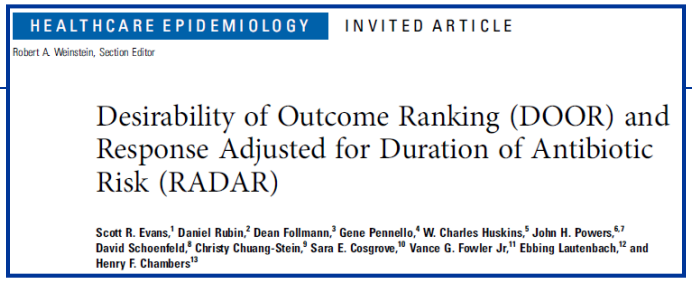
	Today	Tomorrow
Endpoints	**Many**	**Global patient outcome**
Patient / Clinician Preferences	**Limited**	**Incorporated**
Treatment Effects	**One**	**Many (personalized)**

ARLG
Antibacterial Resistance Leadership Group

The Future of Clinical Trials

	Today	Tomorrow
Endpoints	**Many**	**Global patient outcome**
Patient / Clinician Preferences	**Limited**	**Incorporated**
Treatment Effects	**One**	**Many (personalized)**

ARLG
Antibacterial Resistance Leadership Group



HEALTHCARE EPIDEMIOLOGY INVITED ARTICLE

Robert A. Weinstein, Section Editor

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

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

Before we analyze several hundred patients,
we must understand how to analyze one.

- *The patient journey*: "exit examination" or "discharge review" based on a synthesis of benefits, harms, QOL
- DOOR probability: probability of a more desirable global outcome when assigned to the new vs. the control treatment

Example

Motivating question:

Should we use ceftazidime-avibactam or colistin for the initial treatment of CRE infection?

ACCEPTED MANUSCRIPT

Colistin vs. Ceftazidime-avibactam in the Treatment of Infections due to Carbapenem-Resistant Enterobacteriaceae

David van Duin, M.D., Ph.D. ✉, Judith J Lok, Ph.D., Michelle Earley, M.S., Eric Cober, M.D., Sandra S Richter, M.D., Federico Perez, M.D., Robert A Salata, M.D., Robert C Kalayjian, M.D., Richard R Watkins, M.D., M.S., Yohei Doi, M.D., Ph.D. Keith S Kaye, M.D., M.P.H., Vance G Fowler, Jr., M.D., M.H.S., David L Paterson, M.D., Ph.D., Robert A Bonomo, M.D., Scott Evans, Ph.D., for the Antibacterial Resistance Leadership Group.

Clinical Infectious Diseases, cix783, https://doi-org.ezp-prod1.hul.harvard.edu/10.1093/cid/cix783
Published: 04 September 2017 **Article history** ▾

DOOR

- DOOR with 4 levels
  - Alive; discharged home
  - Alive; not discharged home; no renal failure
  - Alive; not discharged home; renal failure
  - Death
- Looking for northward migration of patients in these categories

DOOR

	Colistin (N=46)	Caz-Avi (N=26)
Discharged home	4 (9%)	6 (23%)
Alive; not discharged home; no renal failure	25 (54%)	17 (65%)
Alive; not discharged home; renal failure	5 (11%)	1 (4%)
Death	12 (26%)	2 (8%)

- IPTW-adjusted DOOR Probability: 64% (53%, 75%)
- IPTW-adjusted Win Ratio: 3.0 (1.32, 9.72)

IPTW adjustments: Pitt score, infection type (BSI vs. UTI), and creatinine (sensitivity analyses only)

Challenges

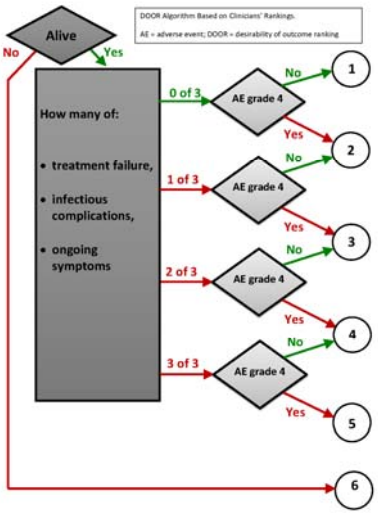
- Cultural change
- Construction of ordinal DOOR is novel and challenging
- Careful deliberation is essential to synthesize the outcomes
- An example strategy …

BAC DOOR

- ARLG conducted a pre-trial sub-study to develop DOOR in *Staphylococcus aureus* bacteremia
- 20 representative patient profiles (benefits, harms, and QoL) constructed based on experiences observed in prior trials
- Profiles sent to 43 expert clinicians. They were asked to rank the patient profiles by desirability of outcome.
- Examined clinician consensus and component outcomes that drive clinician rankings

Decision Tree Algorithm

- Things that we learned
  - Cumulative effect
  - Symptoms important
  - Major non-fatal outcomes had similar importance



ARLG
Antibacterial Resistance Leadership Group

Can we account for:

1. Potential unequal steps between categories?

2. Varying perspectives among patients / clinicians regarding the desirability of the categories?

ARLG
Antibacterial Resistance Leadership Group

The Future of Clinical Trials

	Today	Tomorrow
Endpoints	Many	Global patient outcome
Patient / Clinician Preferences	Limited	Incorporated
Treatment Effects	One	Many (personalized)

PARTIAL CREDIT

	Score
Discharged home	100
Alive; not discharged home; no renal failure	Partial credit
Alive; not discharged home; renal failure	Partial credit
Death	0

Partial Credit: How Much?

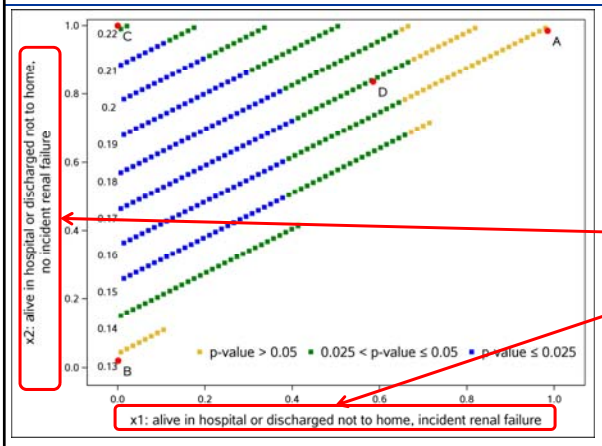
- Transparency and pre-specification important
- Strategies
  - Survey expert clinicians for grading key
  - Patient-guided using QOL

Partial Credit

People have different perspectives.

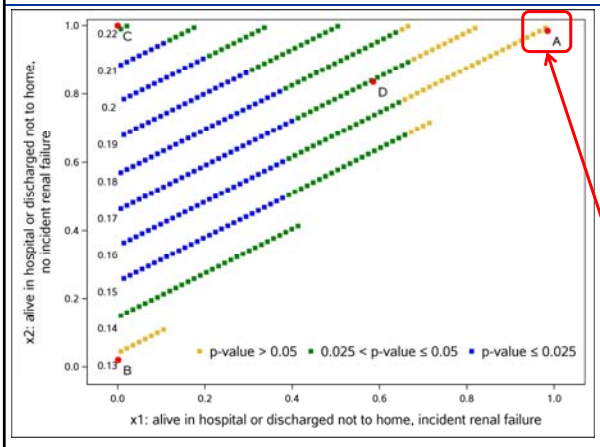
Display treatment contrast as partial credit varies, allowing people to make their own choices based on their own value system.

Contours of Effects as Partial Credit Varies



Category	Credit
Discharged home	100
Alive; Not discharged home; No renal failure	Partial credit
Alive; Not discharged home; Renal failure	Partial credit
Death	0

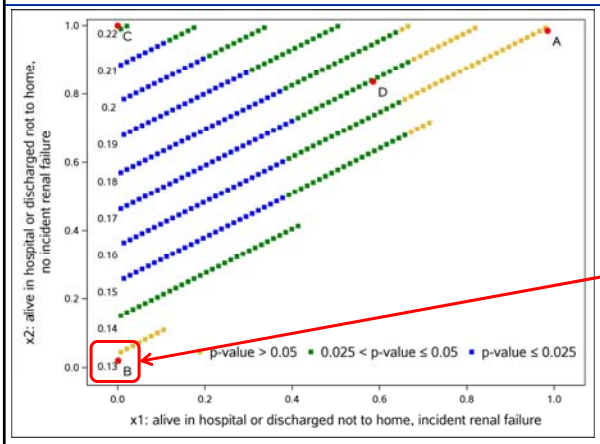
Survival



Category	Credit
Discharged home	100
Alive; Not discharged home; No renal failure	100
Alive; Not discharged home; Renal failure	100
Death	0

Caz-avi advantage: 0.16 (-0.04, 0.32), p = 0.10

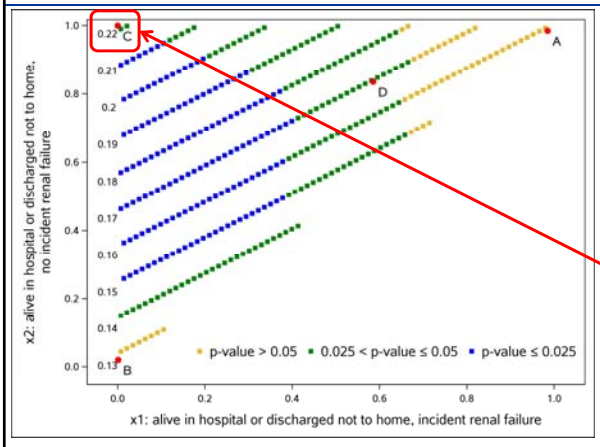
Discharged Home



Category	Credit
Discharged home	100
Alive; Not discharged home; No renal failure	0
Alive; Not discharged home; Renal failure	0
Death	0

Caz-avi advantage: 0.13 (-0.03, 0.31), p = 0.12

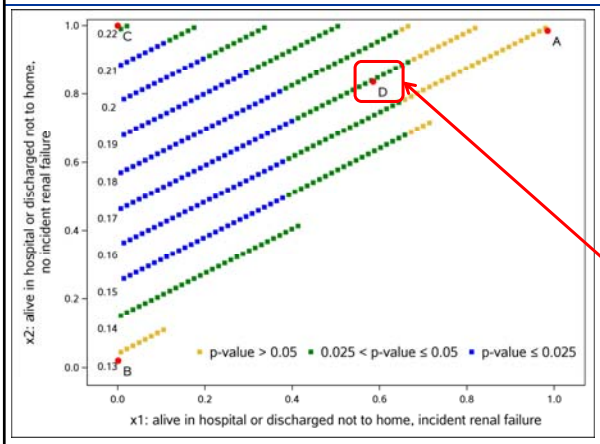
Alive without Renal Failure



Category	Credit
Discharged home	100
Alive; Not discharged home; No renal failure	100
Alive; Not discharged home; Renal failure	0
Death	0

Caz-avi advantage: 0.22 (0.02, 0.40), p = 0.03

Compromise



Category	Credit
Discharged home	100
Alive; Not discharged home; No renal failure	80
Alive; Not discharged home; Renal failure	60
Death	0

Caz-avi advantage: 0.17 (0.01, 0.30), p = 0.04

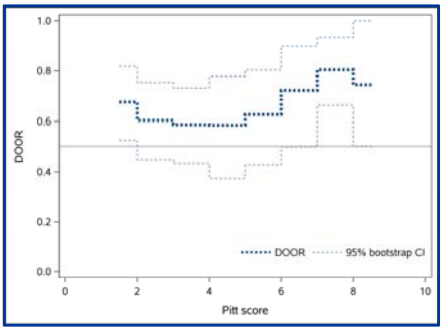
Tailoring Medicine

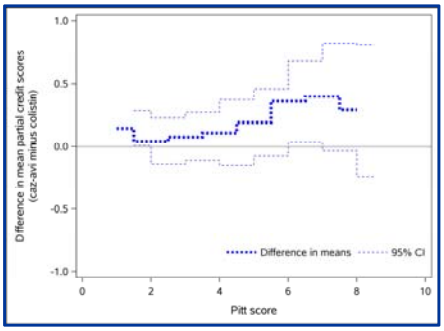
Who benefits from this new therapy?

The Future of Clinical Trials

	Today	Tomorrow
Endpoints	Many	Global patient outcome
Patient / Clinician Preferences	Limited	Incorporated
Treatment Effects	One	Many (personalized)

Caz-Avi-Colistin Contrast as a Function of Disease Severity





DOOR Probability

Partial Credit (80/60)

Largest differences are in the most severe patients.

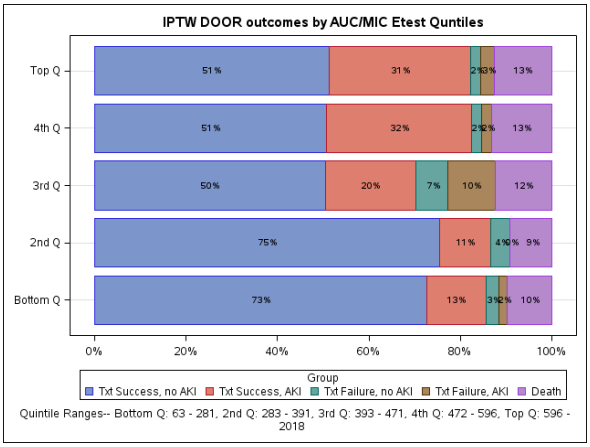
DOOR STEPP

PROVIDE

- Prospective multi-center observational evaluation among adult hospitalized patients with MRSA bloodstream infections
- Research Question
  - What is the vancomycin pharmacodynamic exposure target associated with optimal treatment outcome?
- N=265

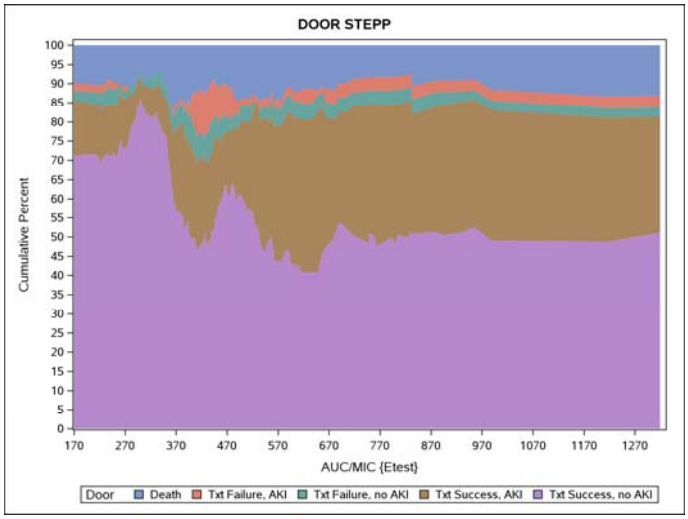
DOOR

DOOR Outcomes by Dosing Quintiles

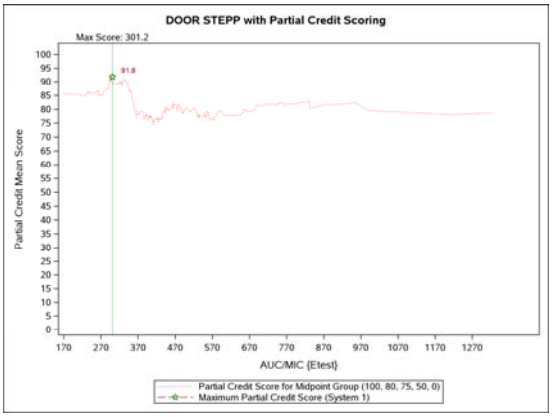


- IPTW adjustments for: presence of infective endocarditis, baseline calculated creatinine clearance, Apache II score, and indicator of any of: prosthetic joint, cardiac prosthetic device, intravascular prosthetic material.

DOOR STEPP



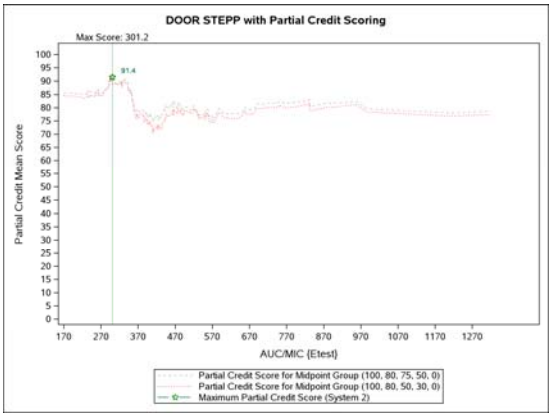
DOOR STEPP: Partial Credit Clinician A



Category	Credit
Treatment Success; No Kidney Injury	100
Treatment Success; Kidney Injury	80
Treatment Failure; No Kidney Injury	75
Treatment Failure; Kidney Injury	50
Death	0

Optimal Dose: 301.2

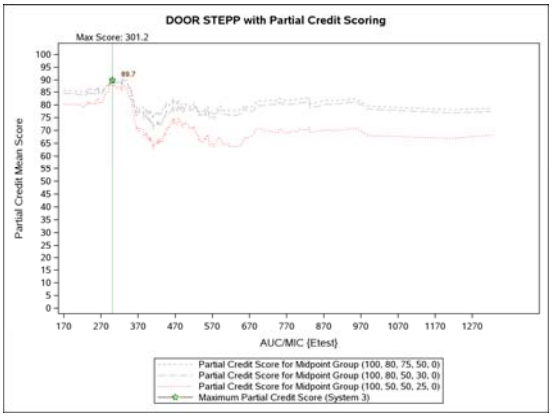
DOOR STEPP: Partial Credit Clinician B



Category	Credit
Treatment Success; No Kidney Injury	100
Treatment Success; Kidney Injury	80
Treatment Failure; No Kidney Injury	50
Treatment Failure; Kidney Injury	30
Death	0

Optimal Dose: 301.2

DOOR STEPP: Partial Credit Clinician C



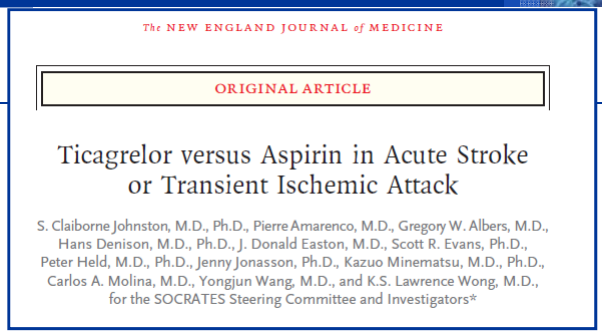
Category	Credit
Treatment Success; No Kidney Injury	100
Treatment Success; Kidney Injury	50
Treatment Failure; No Kidney Injury	50
Treatment Failure; Kidney Injury	25
Death	0

Optimal Dose: 301.2

Cardiovascular Event Prevention Trials?

- Go beyond time-to-first event
  - Include multiple events
- Include major efficacy and safety outcomes (death, stroke, MI, major bleeding)
- Avoid double-counting (fatal bleeding event counted as both a death and a major bleed)
- Avoid competing risks
- Prioritize events by relative importance

ARLG
Antibacterial Resistance Leadership Group



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Ticagrelor versus Aspirin in Acute Stroke or Transient Ischemic Attack

S. Claiborne Johnston, M.D., Ph.D., Pierre Amarenco, M.D., Gregory W. Albers, M.D., Hans Denison, M.D., Ph.D., J. Donald Easton, M.D., Scott R. Evans, Ph.D., Peter Held, M.D., Ph.D., Jenny Jonasson, Ph.D., Kazuo Minematsu, M.D., Ph.D., Carlos A. Molina, M.D., Yongjun Wang, M.D., and K.S. Lawrence Wong, M.D., for the SOCRATES Steering Committee and Investigators*

- International (674 centers) Double-blind RCT of 13,199 patients randomized to Ticagrelor (T) vs. aspirin (ASA) in acute stroke or Transient Ischemic Attach (TIA)
- Primary endpoint: time to stroke, MI, or death within 90 days.
  - 6.7% event rate in T
  - 7.5% event rate in ASA
  - HR=0.89 (0.78, 1.01)

ARLG
Antibacterial Resistance Leadership Group

Five Categories	Treatment A (N=XXXX) Patients (%)	Treatment B (N=XXXX) Patients (%)
Survived with no events		
Survived with 1 non-disabling stroke, MI or PLATO Major Bleed but without experiencing disabling stroke (mRS>=2)		
Survived with >1 non-disabling stroke, MI or PLATO Major Bleed but without experiencing disabling stroke (mRS>=2)		
Survived with disabling stroke (mRS>=2)		
Death		

ARLG
Antibacterial Resistance Leadership Group

Comparing Patient and Physician Risk Tolerance for Bleeding Events Associated with Anticoagulants in Atrial Fibrillation—evidence from the United States and Japan

Ken Okumura, MD[1,*], Hiroshi Inoue, MD, PhD, FACC[2], Masahiro Yasaka, MD, PhD[3], Juan Marcos Gonzalez, PhD[4], A. Brett Hauber[4], Bennett Levitan, MD, PhD[5], Zhong Yuan, MD, PhD[5], Jean-Baptiste Briere, PharmD, MSc[6]

[1]Division of Cardiology, Hirosaki University Graduate School of Medicine, Hirosaki-city, Aomori, Japan; [2]Second Department of Internal Medicine, University of Toyama School of Medicine, Toyama-city, Toyama, Japan; [3]Department of Cerebrovascular Disease, Kyushu Medical Center, Fukuoka-city, Fukuoka, Japan; [4]RTI Health Solutions, Research Triangle Park, NC, USA; [5]Janssen Research & Development LLC, Titusville, NJ, USA; [6]Bayer Pharma AG, Global Health Economics & Outcomes Research Department, Berlin, Germany

VALUE IN HEALTH REGIONAL ISSUES 6C (2015) 65–72

Table 5 – Relative importance estimates of nonfatal cardiovascular events.*

Outcome	Relative importance (SE) of change from 0% to 1% risk of event.†			
	US patients	US physicians	Japanese patients	Japanese physicians
All-cause death	1.00 (.)	1.00 (.)	1.00 (.)	1.00 (.)
Nondisabling stroke	0.75 (0.14)	0.52 (0.06)	0.32 (0.15)	0.13 (0.15)
Disabling, nonfatal stroke	1.68 (0.21)	0.65 (0.07)	0.62 (0.26)	0.16 (0.08)
Myocardial infarction	0.47 (0.05)	0.28 (0.03)	0.13 (0.05)	0.04 (0.05)
Non-CNS, systemic embolism	0.58 (0.09)	0.19 (0.03)	0.14 (0.09)	0.05 (0.05)
Nonmajor clinically relevant bleeding	0.29 (0.04)	0.09 (0.02)	0.04 (0.07)	0.06 (0.02)
Extracranial major bleeding	0.92 (0.15)	0.43 (0.04)	0.24 (0.10)	0.16 (0.06)

CNS, central nervous system; SE, standard error.
* Apart from all-cause death, all events presented in this table are nonfatal events.
† Relative importance is equivalent to preventing deaths per 10,000 patient-years.

ARLG
Antibacterial Resistance Leadership Group

Original article

Relative importance of benefits and risks associated with antithrombotic therapies for acute coronary syndrome: patient and physician perspectives

Zhong Yuan
Bennett Levitan
Department of Epidemiology, Janssen Research & Development, Titusville, NJ, USA
Paul Burton
Clinical Trials Management, Janssen Research & Development, Raritan, NJ, USA
Christine Poulos
A. Brett Hauber
RTI Health Solutions, Research Triangle Park, NC, USA
Jesse A. Berlin
Department of Epidemiology, Johnson & Johnson, New Brunswick, NJ, USA

Current Medical Research & Opinion Vol. 30, No. 9, 2014, 1733–1741

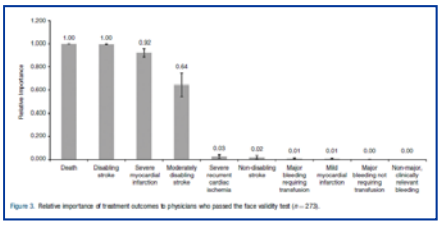


Figure 3. Relative importance of treatment outcomes to physicians who passed the face validity test (n = 273).

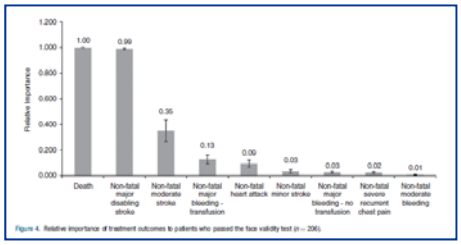


Figure 4. Relative importance of treatment outcomes to patients who passed the face validity test (n = 206).

DIAGNOSTICS

Pragmatic Benefit:risk Evaluation

Diagnostics: Diagnostic Yield

Suppose there is a choice between 2 tests:
one with a higher sensitivity and one with a higher specificity.

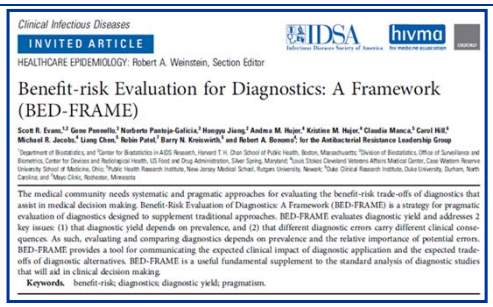
Which test should be selected to optimize clinical outcomes?

How do we evaluate the global utility?

Accuracy (percent correctly classified)

- Range: 0–100%; higher scores indicate better accuracy
- Two challenges to interpretation
  1. Treats all errors as equally important
  2. Not comparable from study to study, since it depends on the prevalence which can vary between studies

BED-FRAME



Clinical Infectious Diseases

INVITED ARTICLE

HEALTHCARE EPIDEMIOLOGY: Robert A. Weinstein, Section Editor

Benefit-risk Evaluation for Diagnostics: A Framework (BED-FRAME)

Scott R. Evans,[1,2] Gene Pennello,[3] Norberto Pantoja-Galicia,[3] Hongyu Jiang,[2] Andrea M. Hujer,[4] Kristine M. Hujer,[4] Claudia Manca,[5] Carol Hill,[6] Michael R. Jacobs,[4] Liang Chen,[3] Robin Patel,[7] Barry N. Kreiswirth,[5] and Robert A. Bonomo[4]; for the Antibacterial Resistance Leadership Group

[1]Department of Biostatistics, and [2]Center for Biostatistics in AIDS Research, Harvard T. H. Chan School of Public Health, Boston, Massachusetts; [3]Division of Biostatistics, Office of Surveillance and Biometrics, Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, Maryland; [4]Louis Stokes Cleveland Veterans Affairs Medical Center, Case Western Reserve University School of Medicine, Ohio; [5]Public Health Research Institute, New Jersey Medical School, Rutgers University, Newark; [6]Duke Clinical Research Institute, Duke University, Durham, North Carolina; and [7]Mayo Clinic, Rochester, Minnesota

The medical community needs systematic and pragmatic approaches for evaluating the benefit-risk trade-offs of diagnostics that assist in medical decision making. Benefit-Risk Evaluation of Diagnostics: A Framework (BED-FRAME) is a strategy for pragmatic evaluation of diagnostics designed to supplement traditional approaches. BED-FRAME evaluates diagnostic yield and addresses 2 key issues: (1) that diagnostic yield depends on prevalence, and (2) that different diagnostic errors carry different clinical consequences. As such, evaluating and comparing diagnostics depends on prevalence and the relative importance of potential errors. BED-FRAME provides a tool for communicating the expected clinical impact of diagnostic application and the expected trade-offs of diagnostic alternatives. BED-FRAME is a useful fundamental supplement to the standard analysis of diagnostic studies that will aid in clinical decision making.

Keywords. benefit-risk; diagnostics; diagnostic yield; pragmatism.

JOURNAL OF BIOPHARMACEUTICAL STATISTICS
2016, VOL. 26, NO. 6, 1083–1097
http://dx.doi.org/10.1080/10543406.2016.1226335

Taylor & Francis
Taylor & Francis Group

Comparing diagnostic tests on benefit-risk

Gene Pennello[a], Norberto Pantoja-Galicia[a], and Scott Evans[b]

[a]Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, Maryland, USA; [b]Center for Biostatistics in AIDS Research and the Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA

Weighted Accuracy

- Accuracy adjusted for the relative importance (r) of a false positive vs. a false negative
- Range: 0–100%; higher scores indicate better accuracy

$$WA(p,r) = [rp(specificity) + (1-p)(sensitivity)]\,(1/(rp+1-p))$$

- Depends on the prevalence (can vary by region and timeframe)
  - Use average weighted accuracy (AWA) over the relevant prevalence range

Average Weighted Accuracy (AWA)

Given the relevant prevalence range $p \in [a,b]$,

- $AWA = \frac{1}{b-a}\int_a^b WA(p;r)dp = c_1\,(sensitivity) + c_2(specificty).$

  $c_1 = \frac{1}{1-r} - \frac{r}{(b-a)(1-r)^2} ln\left(\frac{r+b(1-r)}{r+a(1-r)}\right),$

  $c_2 = \frac{r}{(1-r)(b-a)} ln\left(\frac{r+b(1-r)}{r+a(1-r)}\right) - \frac{r}{1-r} + \frac{r^2}{(1-r)^2(b-a)} ln\left(\frac{r+b(1-r)}{r+a(1-r)}\right),$

- $SE\left(\widehat{AWA}\right) = \sqrt{\frac{{c_1}^2\widehat{PPA}(1-\widehat{PPA})}{n_1} + \frac{{c_2}^2\widehat{NPA}(1-\widehat{NPA})}{n_2}}.$

Evaluating Utility Using AWA

- Components
  - Sensitivity and specificity
  - Relevant prevalence range
  - Relative importance of a false positive vs. false negative
  - The goal (the null hypothesis)
- Rigorously evaluate each and pre-specify

Example

- Prospective study evaluating a diagnostic test categorizing acute respiratory tract illness into bacterial vs. non-bacterial etiologies
- Prevalence of bacterial etiology: $p \in [0.1, 0.3]$

Relative Importance (r)

- Failing to identify bacterial etiology may be a more important error than failing to identify non-bacterial etiology
  - Failing to identify non-bacterial etiology
    → Unnecessary exposure to antibiotics
  - Failing to identify bacterial etiology
    → Failure to treat with necessary antibiotics
- Ideally would have structured evaluation involving comparison of the clinical consequences resulting from treatment decisions based on correct vs incorrect diagnoses in both bacterial and non-bacterial disease though such data may be sparse

Relative Importance (r)

Alternative: obtain a data-driven answer via a survey of experts

"How important is a false bacterial call vs. a false non-bacterial?"

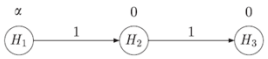
100% implies that the two errors are equally important.

What is the goal?

- Evaluate if AWA for the new test is better than the AWA for:
  - Best random test (BRT)
    - The random test (a test that claims bacterial etiology with a fixed probability) with the highest AWA
  - Diagnostic alternatives
    - Procalcitonin

Null Hypotheses

- A three-step sequential hypothesis testing strategy

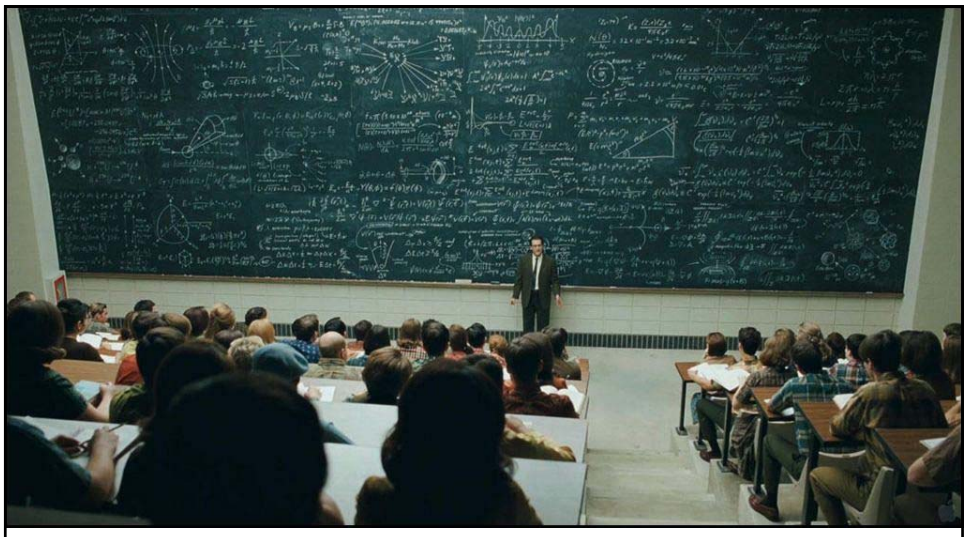


- Hypotheses
  - H_1: One-sample test vs. the BRT
  - H_2: One-sample test vs. the point estimate of the procalcitonin AWA
  - H_3: Two-sample test vs. procalcitonin

ARLG
Antibacterial Resistance Leadership Group

Significant Contributors (p<0.001)

- Dean Follmann
- Dan Rubin
- Gene Pennello
- Norberto Pantoja-Galicia
- Ying Liu
- Judith Lok
- Michelle Earley
- Thuy Tran
- Hongyu Jiang
- Matt Finnemeyer
- Chip Chambers
- David van Duin
- David Patterson
- Tom Holland
- Ephraim Tsalik
- Robert Bonomo
- Barry Kreiswirth



I have no doubt that you will enthusiastically applaud now …
because you are so relieved that it is over.

Thank you.