Desirability Of Outcome Ranking (DOOR)

Scott Evans, PhD, MS

Director, The Biostatistics Center
Founding Chair and Professor, Department of Biostatistics and Bioinformatics
Milken Institute School of Public Health
George Washington University

Disclosures

- Grants: from NIAID, NCI of the NIH.
- Royalties: Chapman & Hall/CRC Press.
- Board Membership: Frontier Science Foundation
- Advisory Committee and DSMB service: FDA, NIH, Analgesic Anesthetic and Addiction Clinical Trial Translations Innovations Opportunities and Networks (ACTTION), Pfizer, Roche, AstraZeneca, Vir, GSK, Genentech, Johnson & Johnson, International Drug Development Institute, Breast International Group, University of Pennsylvania, Duke University, Wake Forest University, Akouous, Apellis, Teva, DayOneBio, Rakuten, Abbvie, Nuvelution, Clover, FHI Clinical, Lung Biotech, SAB Biopharm, Candel, Takeda, Rakuten, Eli Lilly, Novartis

Complexities in NASH Clinical Trials

- Several important multi-organ clinical and other events
  - Death
  - CV
  - CKD
  - New onset diabetes
  - Weight
  - ALT
  - Cancer
- Critical to recognize for clinical decision-making and patients, that some events are more important than others
  - Death is more important than a non-fatal event
  - Events w/ disabling sequelae are more important than those w/ non-disabling sequelae
  - Events w/ permanent sequelae are more important than those w/ transient sequelae

Totality of Evidence and the Challenges in Benefit:risk Evaluation

- Typical benefit:risk analyses
  - Compare interventions for each efficacy and safety outcome
  - Combine these effects

- These analyses
  - Fail to incorporate associations between outcomes
  - Fail to recognize the cumulative nature of outcomes on individual patients
  - Suffer from competing risk complexities during interpretation of individual outcomes, and
  - Since efficacy and safety analyses are often conducted on different populations, generalizability is unclear.

Question 1

- 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 2

- 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
- Why do we analyze them separately?

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%

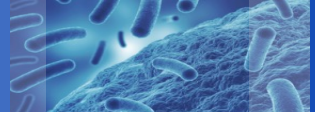
Which treatment would you choose?

They all have the same success rate.

A has the lowest safety event rate.

B and C are indistinguishable.

Choose A…right?



Our culture is to use patients to analyze the outcomes.

Shouldn’t we use outcomes to analyze the patients?

Analysis of Patients: 4 Possible Outcomes

A (N=100)

Success: 50%

Safety event: 30%

		Success +	Success -
SE	+	15	15
	-	35	35

B (N=100)

Success: 50%

Safety event: 50%

	Success +	Success -
+	50	0
-	0	50

C (N=100)

Success: 50%

Safety event: 50%

	Success +	Success -
+	0	50
-	50	0

Analysis of Patients: 4 Possible Outcomes

A (N=100)

Success: 50%

Safety event: 30%

		Success +	Success -
SE	+	15	15
	-	35	35

B (N=100)

Success: 50%

Safety event: 50%

	Success +	Success -
+	50	0
-	0	50

C (N=100)

Success: 50%

Safety event: 50%

	Success +	Success -
+	0	50
-	50	0

Analysis of Patients: 4 Possible Outcomes

A (N=100)

Success: 50%

Safety event: 30%

		Success +	Success -
SE	+	15	15
	-	35	35

B (N=100)

Success: 50%

Safety event: 50%

		Success +	Success -
SE	+	50	0
	-	0	50

C (N=100)

Success: 50%

Safety event: 50%

		Success +	Success -
SE	+	0	50
	-	50	0

Analysis of Patients: 4 Possible Outcomes

A (N=100)

Success: 50%

Safety event: 30%

		Success +	Success -
SE	+	15	15
	-	35	35

B (N=100)

Success: 50%

Safety event: 50%

		Success +	Success -
SE	+	50	0
	-	0	50

C (N=100)

Success: 50%

Safety event: 50%

		Success +	Success -
SE	+	0	50
	-	50	0

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.

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

"The order of operations is important…"

Desirability Of Outcome Ranking (DOOR)

- A paradigm for the design, monitoring, analyses and reporting of clinical trials based on patient centric benefit:risk

- Addresses noted challenges

Before we analyze several hundred patients,

we must understand how to analyze one.

Brief Outline of DOOR

- Use outcomes to analyze patients
  - Construct ordinal DOOR based on the *patient journey*

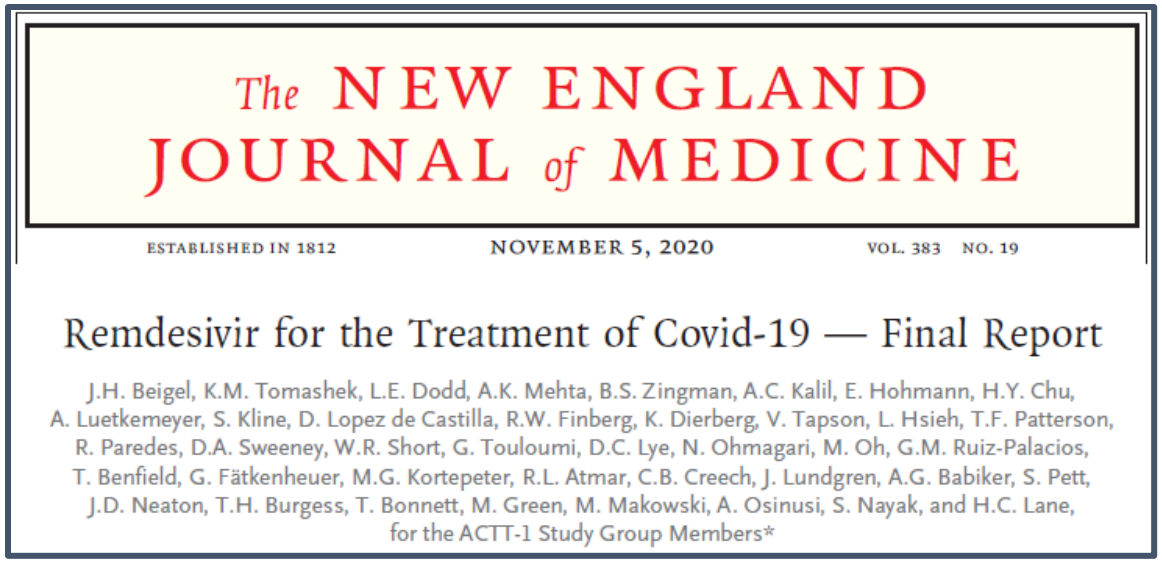
- Two complimentary analyses
  1. Rank-based
     - Estimating the DOOR probability: the probability that a patient from treatment has a more desirable outcome than a patient on control
       - 50% implies equivalence
       - Intuitively attractive
  2. Partial credit (score based analyses)

- Analyze individual outcomes for comprehensive assessment

Adaptive Covid-19 Treatment Trial (ACTT-1)

- No known efficacious treatments for COVID-19 at the time

- ACTT-1
  - Randomized double-blind placebo-controlled trial of IV remdesivir in hospitalized adult COVID-19 patients w/ LRTI
  - N=1062



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 NOVEMBER 5, 2020 VOL. 383 NO. 19

Remdesivir for the Treatment of Covid-19 — Final Report

J.H. Beigel, K.M. Tomashek, L.E. Dodd, A.K. Mehta, B.S. Zingman, A.C. Kalil, E. Hohmann, H.Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R.W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T.F. Patterson, R. Paredes, D.A. Sweeney, W.R. Short, G. Touloumi, D.C. Lye, N. Ohmagari, M. Oh, G.M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M.G. Kortepeter, R.L. Atmar, C.B. Creech, J. Lundgren, A.G. Babiker, S. Pett, J.D. Neaton, T.H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, and H.C. Lane, for the ACTT-1 Study Group Members*

ACTT-1

- Important events
  - Death
  - Hospitalized with invasive mechanical ventilation / ECMO
  - SAE that is not resolved or resolved with sequelae

	Treatment	
DOOR (Day 29)	**Remdesivir (N=541)**	**Placebo (N=521)**
1. Alive: 0 of the other events above		
2. Alive: 1 of the other events above		
3. Alive: both of the other events above		
4. Death		

ACTT-1

- Important events
  - Death
  - Hospitalized with invasive mechanical ventilation / ECMO
  - SAE that is not resolved or resolved with sequelae

	Treatment	
DOOR (Day 29)	**Remdesivir (N=541)**	**Placebo (N=521)**
1. Alive: 0 of the other events above		382 (73.3%)
2. Alive: 1 of the other events above		57 (10.9%)
3. Alive: both of the other events above		6 (1.2%)
4. Death		76 (14.6%)

ACTT-1

- Important events
  - Death
  - Hospitalized with invasive mechanical ventilation / ECMO
  - SAE that is not resolved or resolved with sequelae

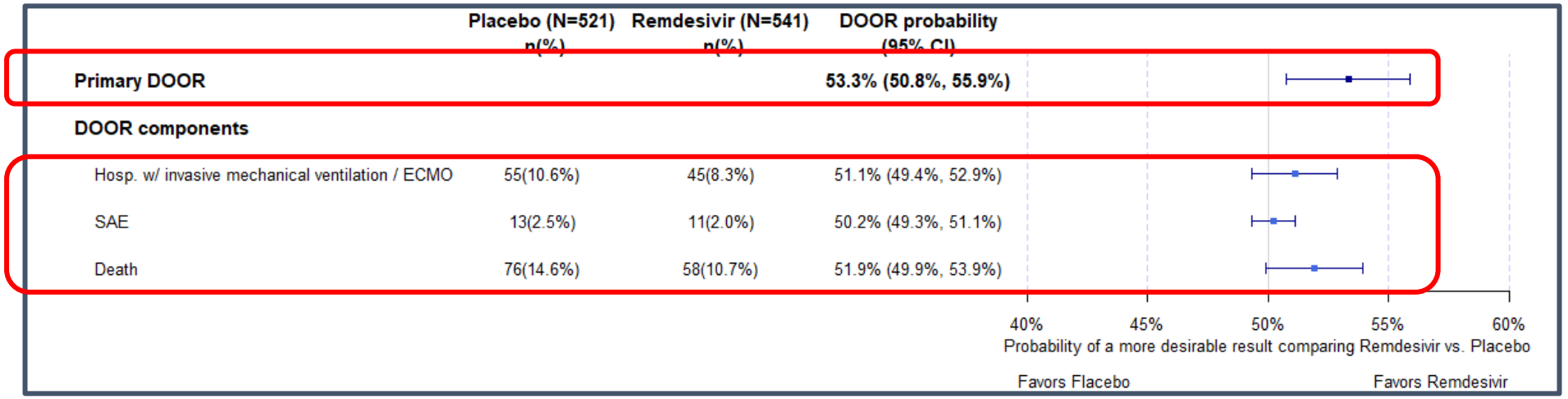
	Treatment	
DOOR (Day 29)	**Remdesivir (N=541)**	**Placebo (N=521)**
1. Alive: 0 of the other events above	433 (80.0%)	382 (73.3%)
2. Alive: 1 of the other events above	42 (7.8%)	57 (10.9%)
3. Alive: both of the other events above	8 (1.5%)	6 (1.2%)
4. Death	58 (10.7%)	76 (14.6%)

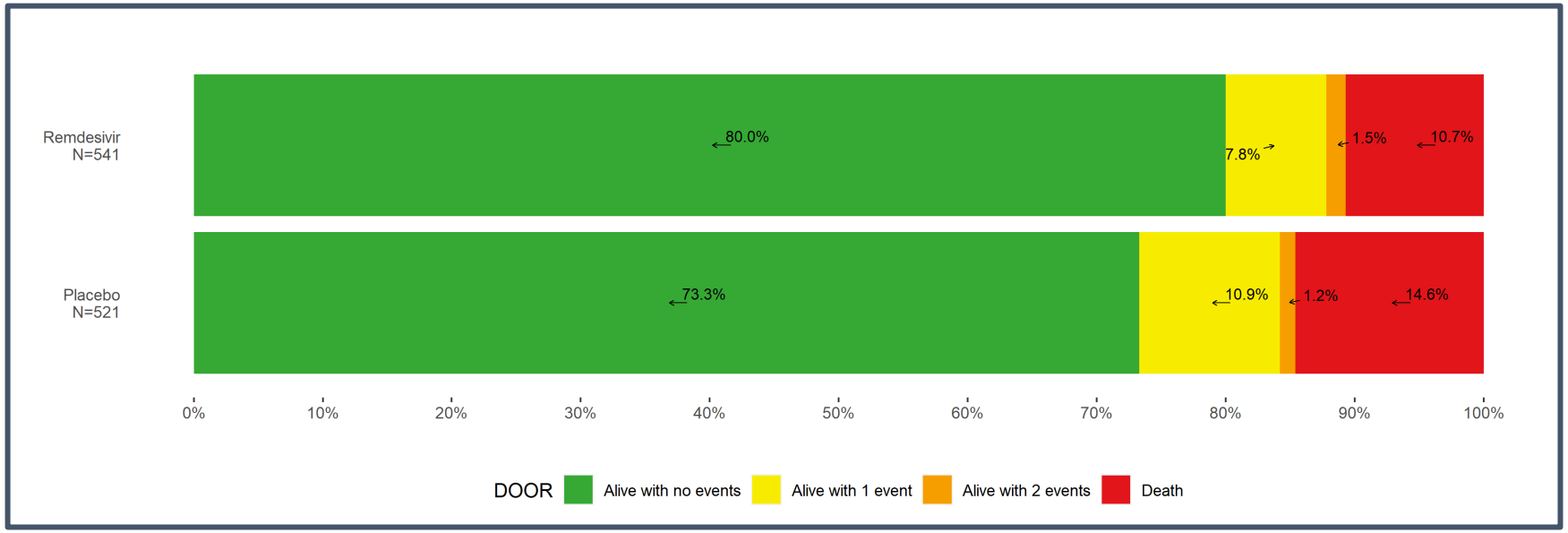
ACTT-1

- Important events
  - Death
  - Hospitalized with invasive mechanical ventilation / ECMO
  - SAE that is not resolved or resolved with sequelae

	Treatment	
DOOR (Day 29)	**Remdesivir (N=541)**	**Placebo (N=521)**
1. Alive: 0 of the other events above	433 (80.0%)	382 (73.3%)
2. Alive: 1 of the other events above	42 (7.8%)	57 (10.9%)
3. Alive: both of the other events above	8 (1.5%)	6 (1.2%)
4. Death	58 (10.7%)	76 (14.6%)

	Placebo (N=521) n(%)	Remdesivir (N=541) n(%)	DOOR probability (95% CI)
Primary DOOR			**53.3% (50.8%, 55.9%)**
DOOR components			
Hosp. w/ invasive mechanical ventilation / ECMO	55(10.6%)	45(8.3%)	51.1% (49.4%, 52.9%)
SAE	13(2.5%)	11(2.0%)	50.2% (49.3%, 51.1%)
Death	76(14.6%)	58(10.7%)	51.9% (49.9%, 53.9%)





Expected Gain/Loss when Treating 1000 Patients

DOOR	Placebo	Remdesivir	Gained (+) / prevented (-) with treatment	Cumulative gained (+) / prevented (-) with treatment
1: Alive with no events	733	800	67	67
2: Alive with one event	109	78	-31	36
3: Alive with both events	12	15	3	39
4: Death	146	107	-39	0

Components	Placebo	Remdesivir	Gained(+)/prevented(-) with treatment
Hospitalized with invasive mechanical ventilation / ECMO	106	83	-23
SAE	25	20	-5
Death	146	107	-39

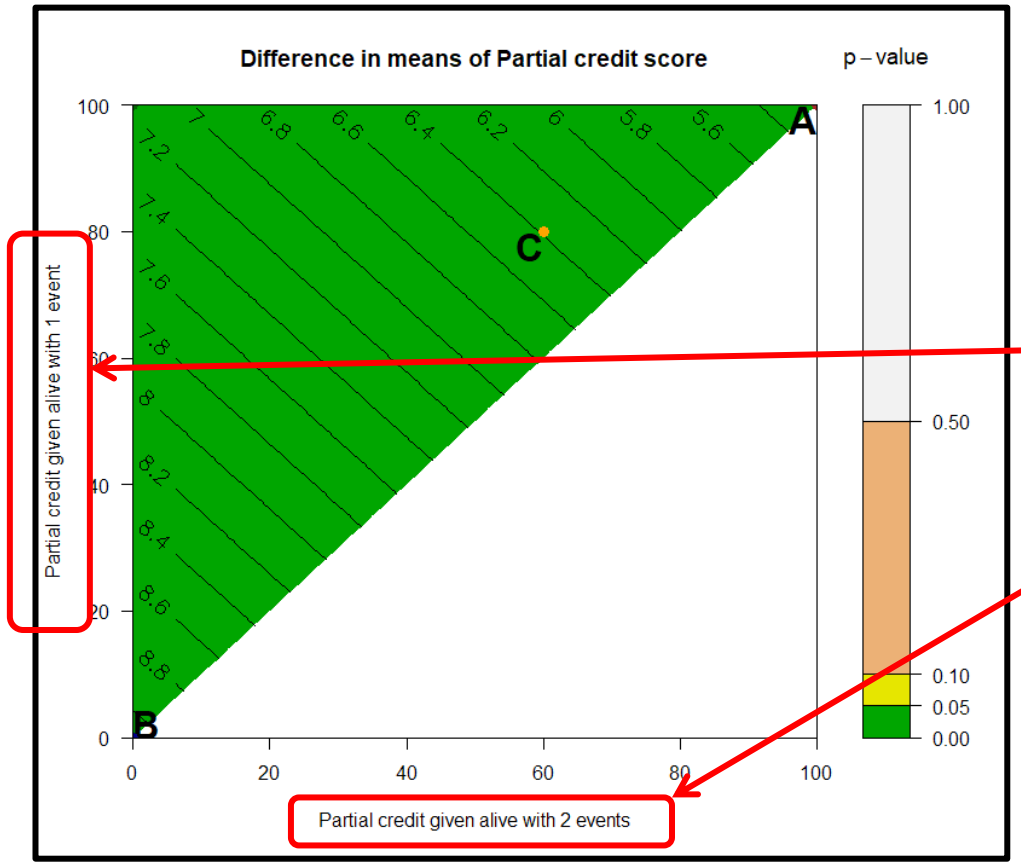
PARTIAL CREDIT

	Score
1. Alive: 0 of the events	100
2. Alive: 1 of the events	Partial credit
3. Alive: both of the events	Partial credit
4. Death	0

Partial credit can be used to account for:

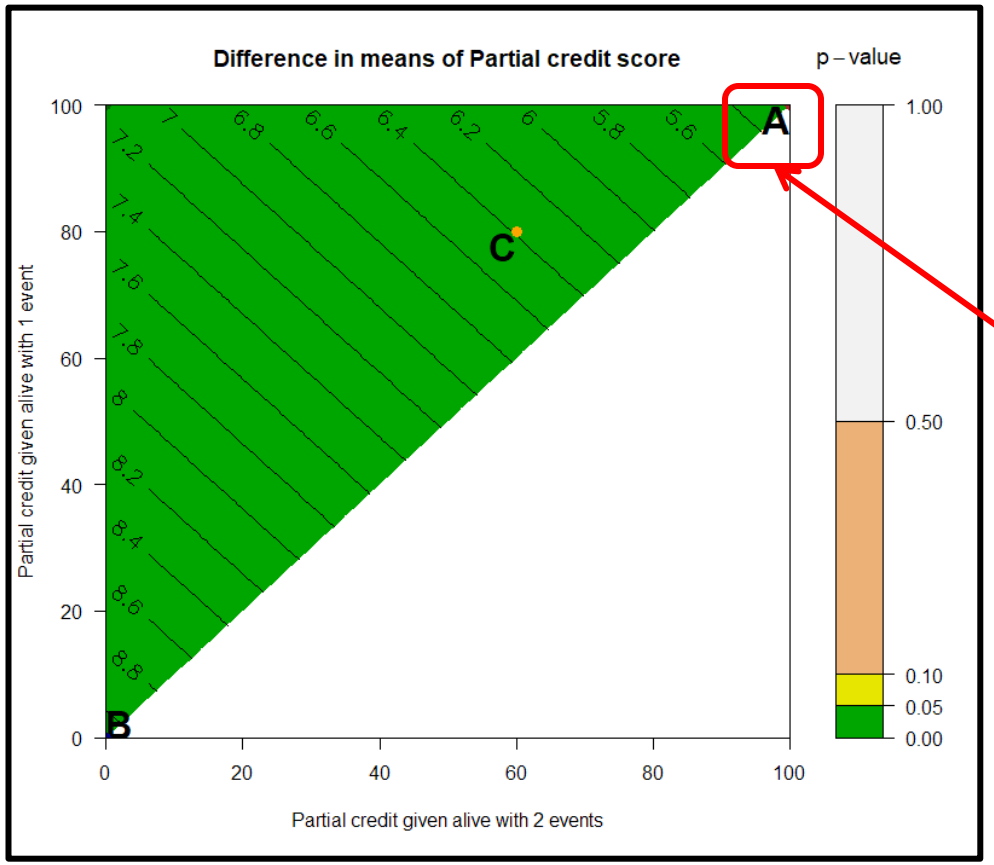
1. **Unequal steps between categories**
2. **Personalized perspectives among patients / clinicians regarding the desirability of the categories**
3. **Robustness analyses**

Contours of Effects as Partial Credit Varies



Category	Credit
Alive; 0 event	100
Alive; 1 event	Partial credit
Alive; both events	Partial credit
Death	0

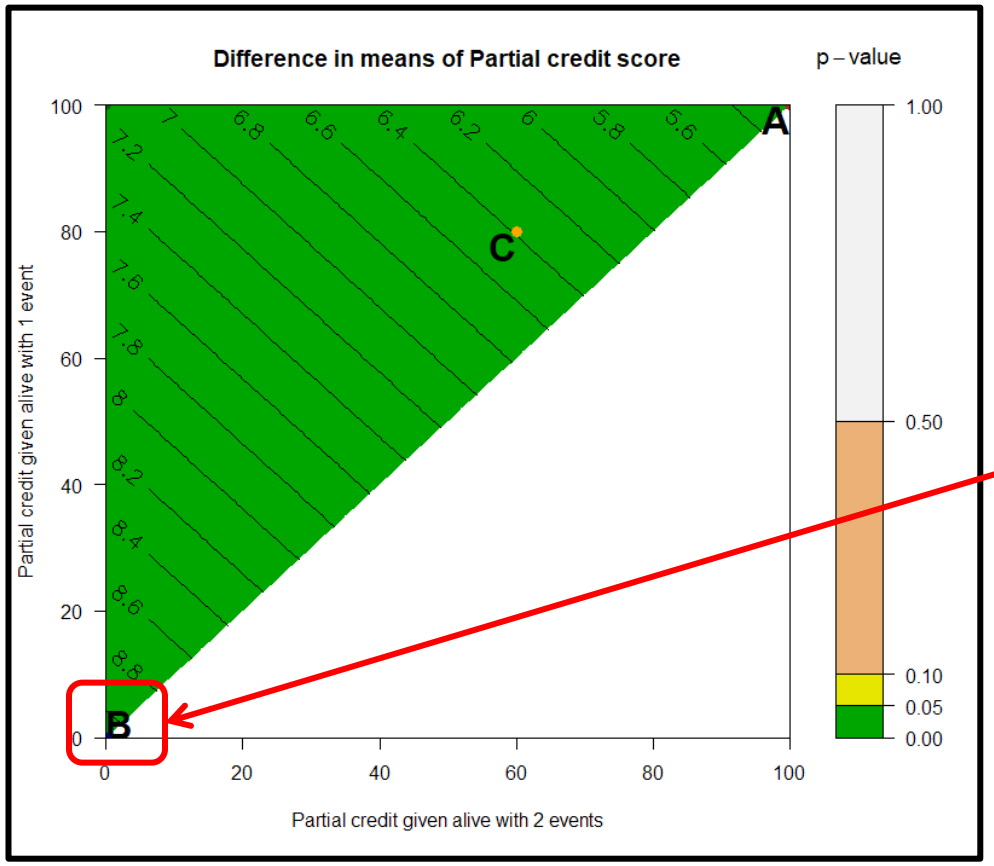
Survival



Category	Credit
Alive 0 events	100
Alive; 1 event	100
Alive; both events	100
Death	0

Remdesivir Advantage ≈ 5.2%

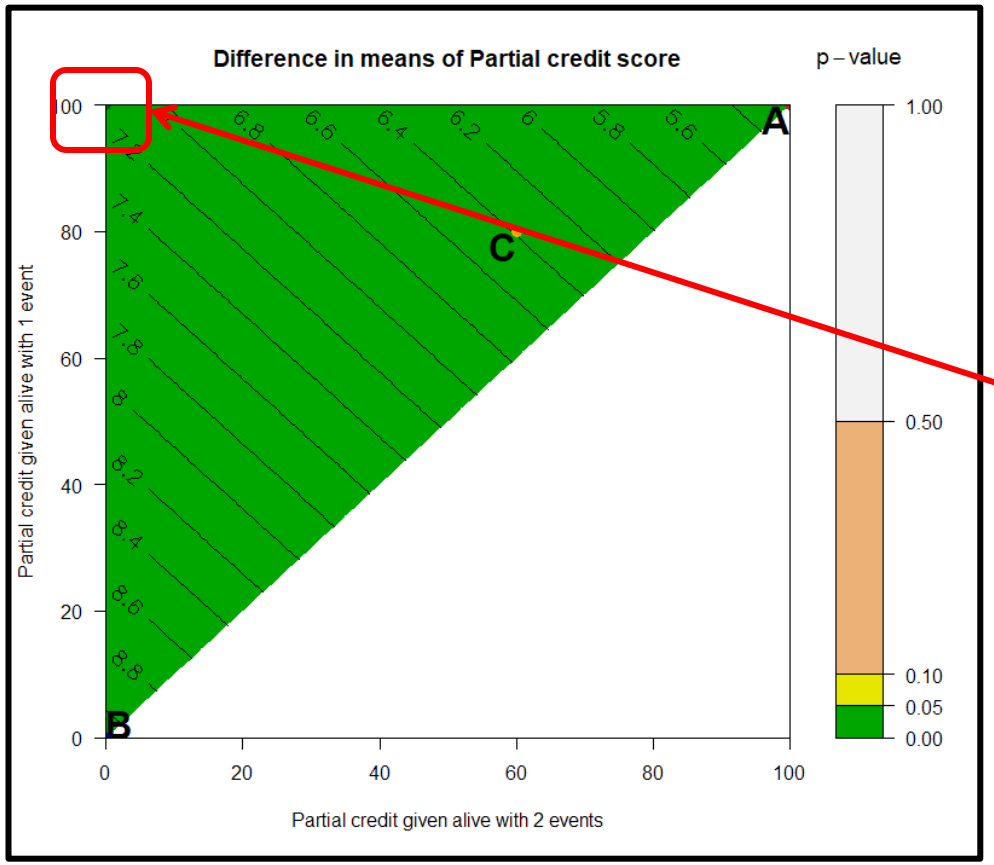
Alive; 0 Events



Category	Credit
Alive; 0 events	100
Alive; 1 event	0
Alive; both events	0
Death	0

Remdesivir Advantage ≈ 9.0%

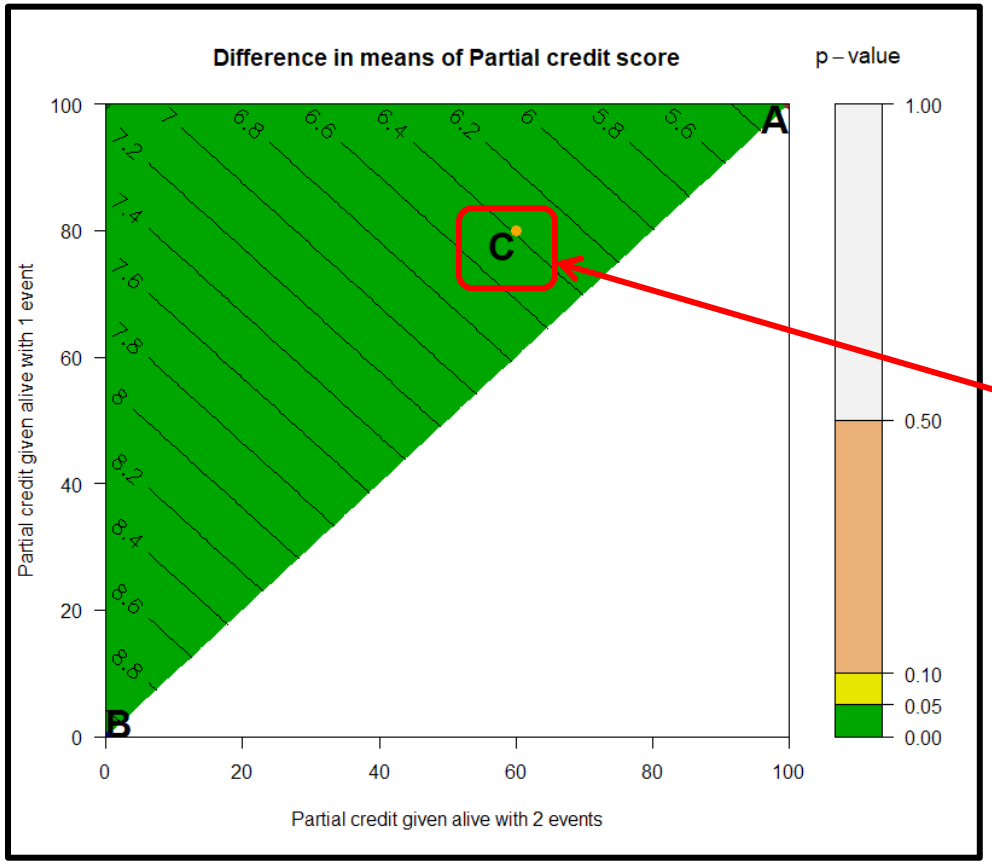
Alive with 0 or 1 Events



Category	Credit
Alive; 0 events	100
Alive; 1 event	100
Alive; both events	0
Death	0

Remdesivir Advantage ≈ 7.2%

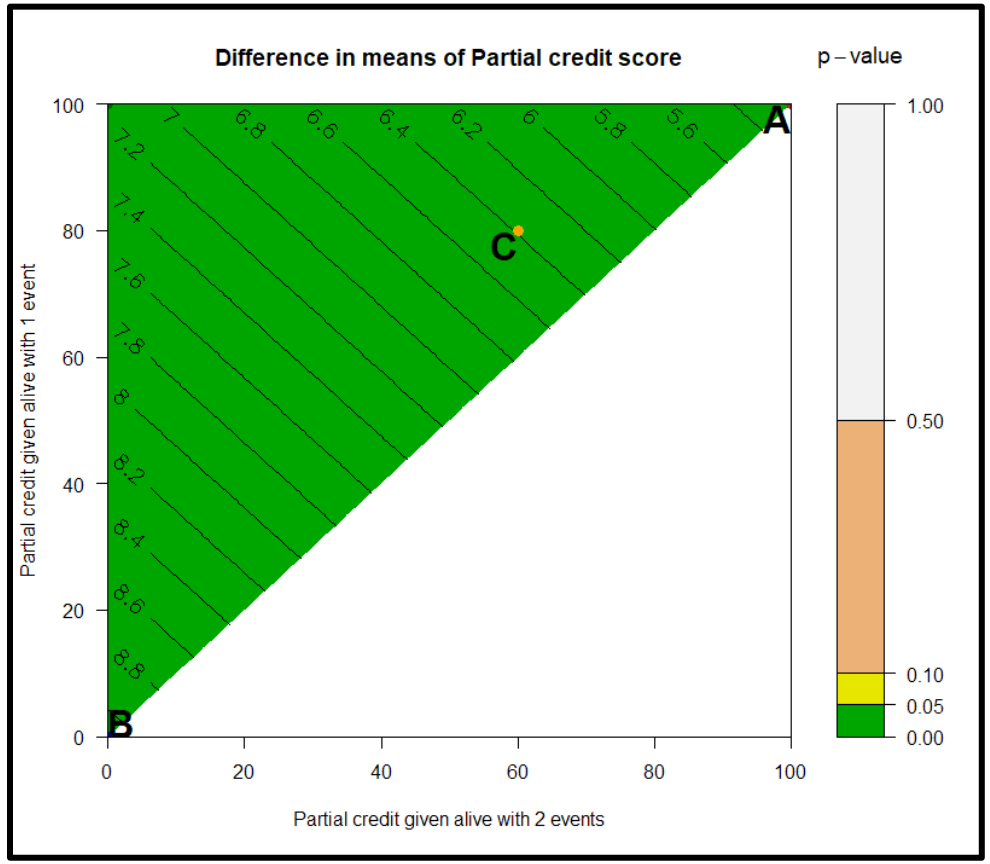
Compromise



Category	Credit
Alive; 0 events	100
Alive; 1 event	80
Alive; both events	60
Death	0

Remdesivir Advantage ≈ 6.4%

Robustness

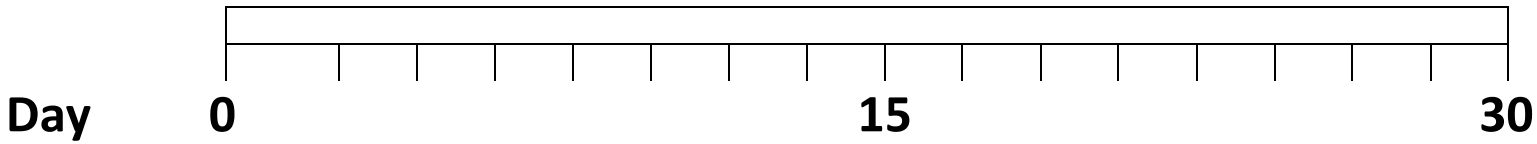


- Numeric results vary by partial credit grading key, though robustness is demonstrated as green color indicates statistical significance everywhere

Anthology of Patient Stories

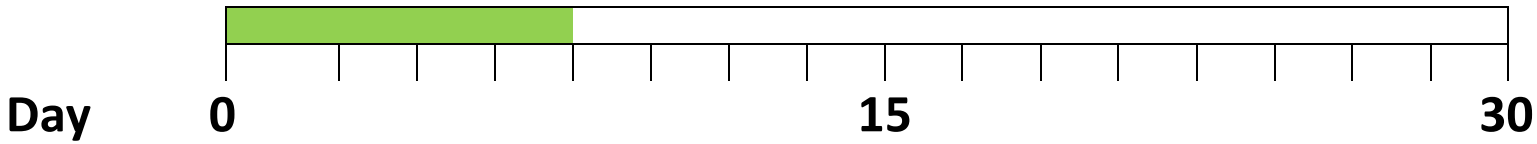
Anthology of Patient Stories

The Patient Story



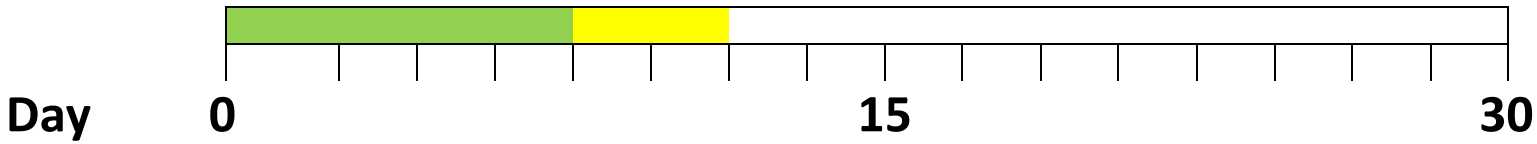
Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story



Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story



Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story

Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story

Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story

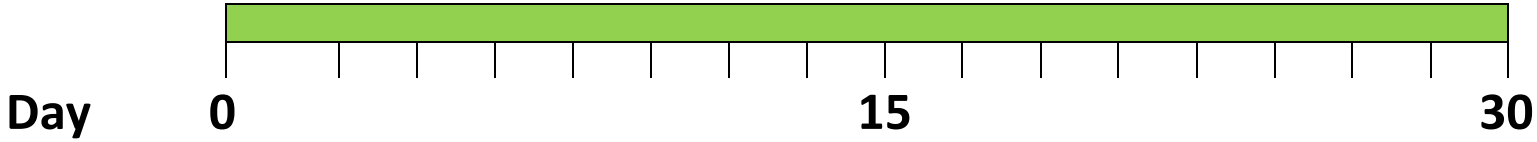
Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Patient Story

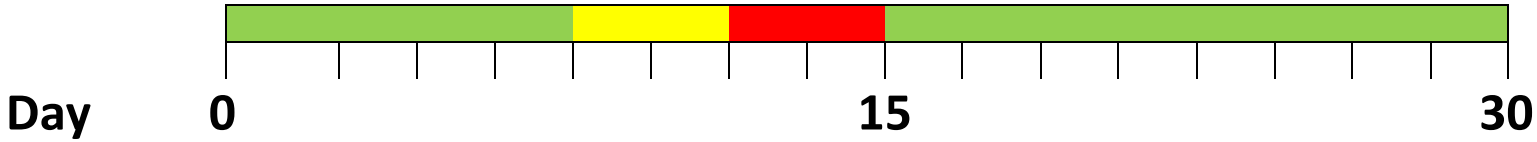
Alive with 0 events	
Alive with 1 event	
Alive with 2 events	
Death	

The Trial Anthology: A Collection of Patient Stories

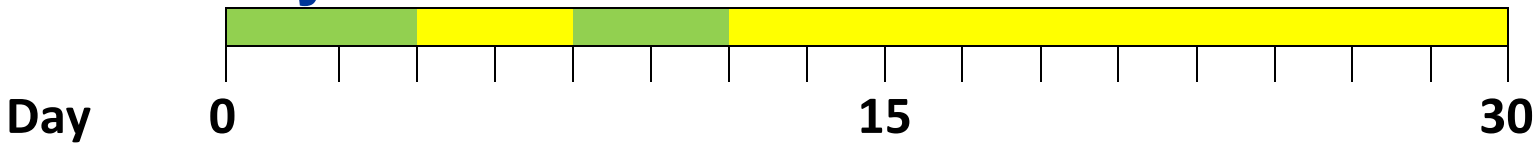
Saul Goodman



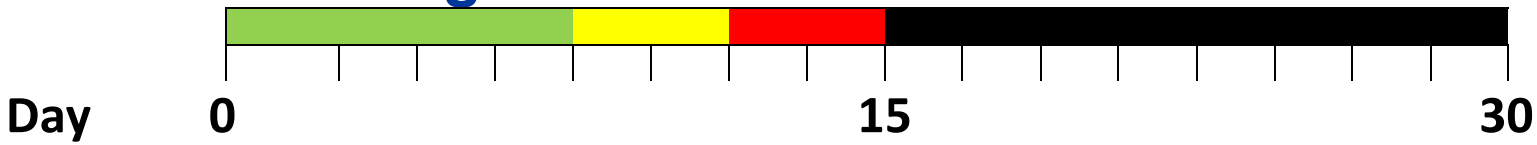
Ita Lendswell



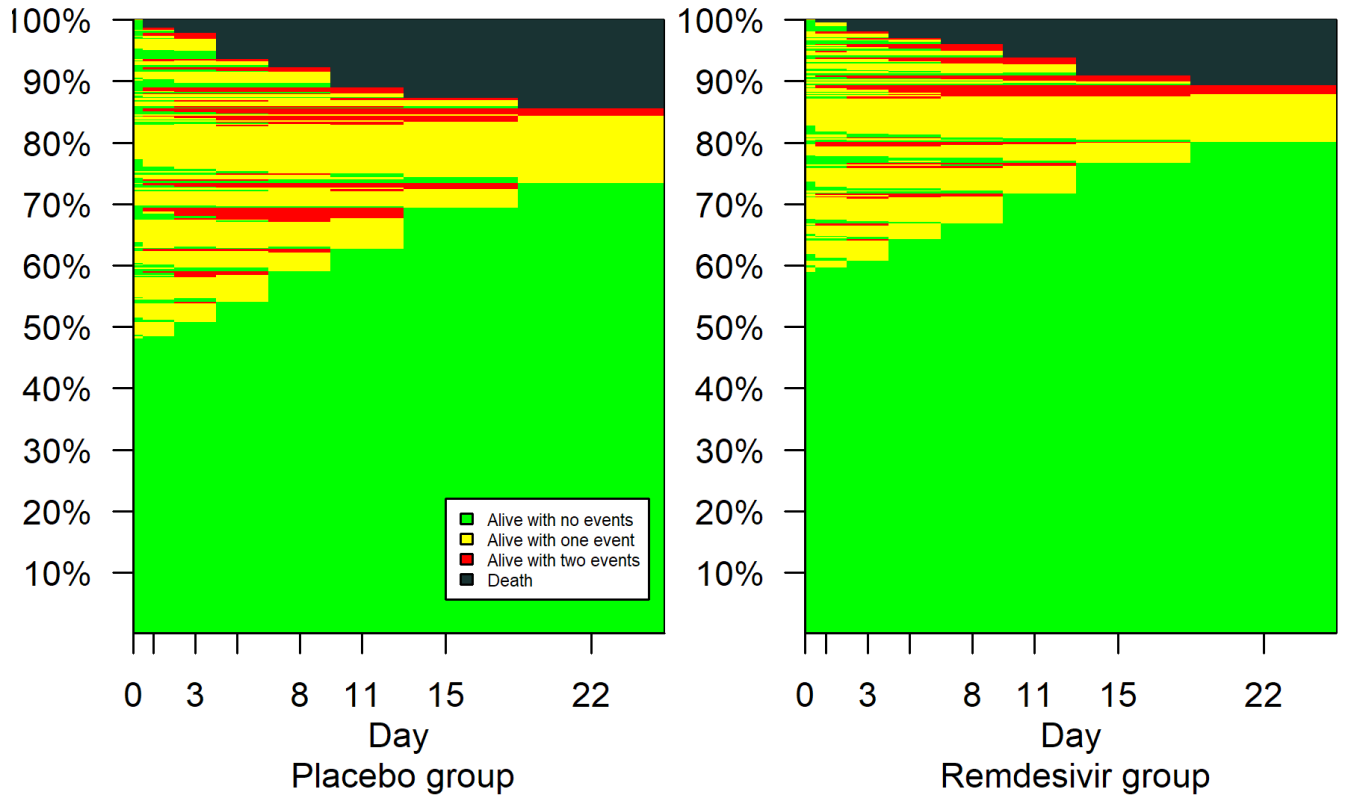
Nori Kovery



Statistician Marge N. O’vera



The Trial Anthology of Patient Stories



- Mortality at Day 29: 14.6% in placebo; 10.7% in Remdesivir
- No events at Day 29: 73.3% in placebo; 80% in Remdesivir
- No events in all time intervals: 48% in placebo; 58.8% in Remdesivir

- Council for International Organizations of Medical Sciences (CIOMS) in Geneva is expected to recommend regular including of DOOR in trial protocols to enhance benefit:risk assessment

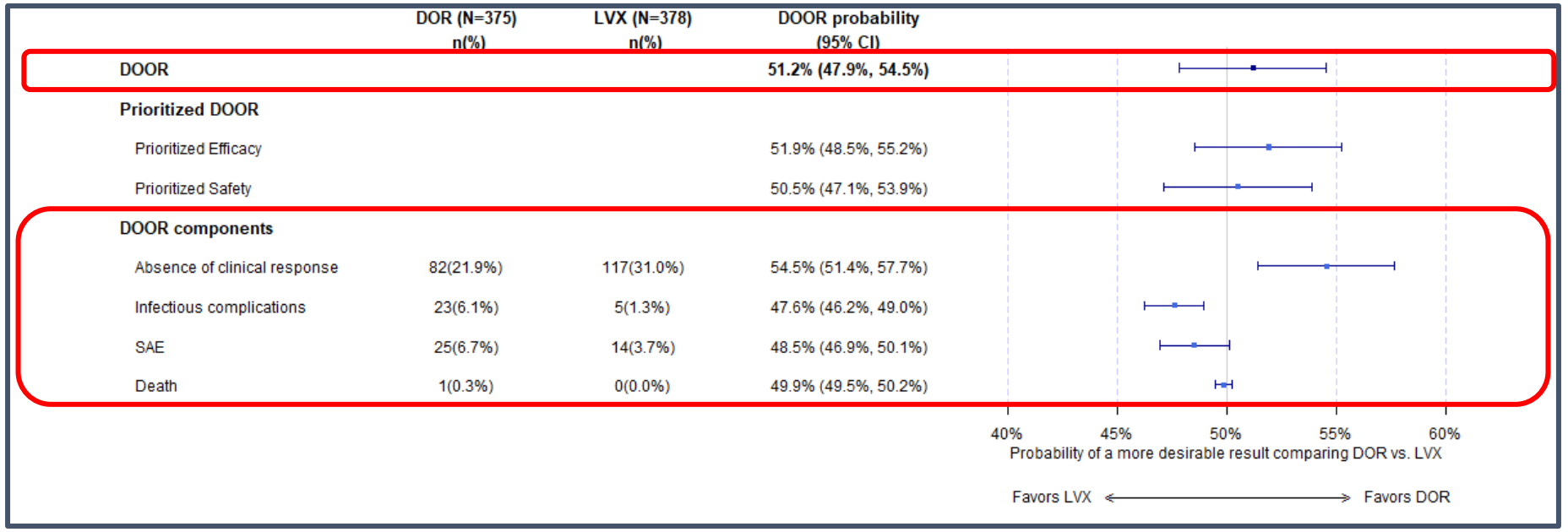
ARLG Innovations Working Group

- Group of Antibacterial Resistance Leadership Group (ARLG) investigators and regulators
- Goal: develop a standardized DOOR outcome and approach for each regulatory indication associated with bacterial infections
- Industry provides data from completed trials to refine and apply the methods
- FDA conducts analyses on data that they have to inform development

Regulatory Indications for Bacterial Infections

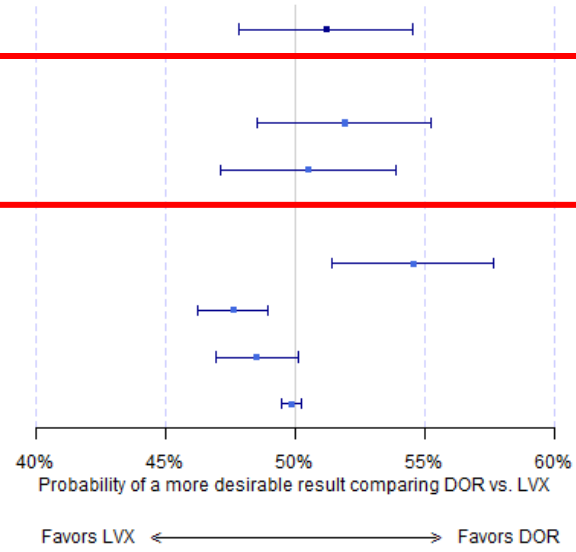
	HABP/VABP	ABSSSI	cUTI	cIAI
Absence of clinical response *(Lack of global resolution of index infection or recurrence of index infection before test of cure*	• Did not meet clinical success or cure as assessed by study investigator at test of cure • Recurrent HABP/VABP prior to test of cure	• Did not meet clinical success or cure as assessed by study investigator at test of cure • Recurrent ABSSSI prior to test of cure	• Did not meet clinical success or cure as assessed by study investigator at test of cure • Recurrent cUTI prior to test of cure	• Did not meet clinical success or cure as assessed by study investigator at test of cure • Recurrent cIAI prior to test of cure
Infectious complications *(Newly identified complications or progression of the original infection that was not present at enrollment or C. difficile)*	• Complicated pleural effusion • Lung abscess/necrotizing pneumonia • ARDS • Meningitis • Bacteremia • Septic shock • Need for intubation • *C. difficile*	• Unplanned surgical intervention for progression/complication of original infection • Bacteremia • Septic shock • Osteomyelitis • *C. difficile*	• Renal or intraabdominal abscess • Septic shock • Bacteremia • Recurrent cUTI after test of cure • Prostatic abscess • Epididymo-orchitis • *C. difficile*	• Bacteremia • Septic shock • Peritonitis • Unplanned surgical intervention • *C. difficile*
Serious adverse events (AEs)	• ICH E6 Good Clinical Practice guidelines	• ICH E6 Good Clinical Practice guidelines	• ICH E6 Good Clinical Practice guidelines	• ICH E6 Good Clinical Practice guidelines

Example cUTI trial:
Benefit – Harm Tradeoff



Example cUTI trial:
Benefit – Harm Tradeoff

	DOR (N=375) n(%)	LVX (N=378) n(%)	DOOR probability (95% CI)
DOOR			**51.2% (47.9%, 54.5%)**
Prioritized DOOR			
Prioritized Efficacy			51.9% (48.5%, 55.2%)
Prioritized Safety			50.5% (47.1%, 53.9%)
DOOR components			
Absence of clinical response	82(21.9%)	117(31.0%)	54.5% (51.4%, 57.7%)
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%)



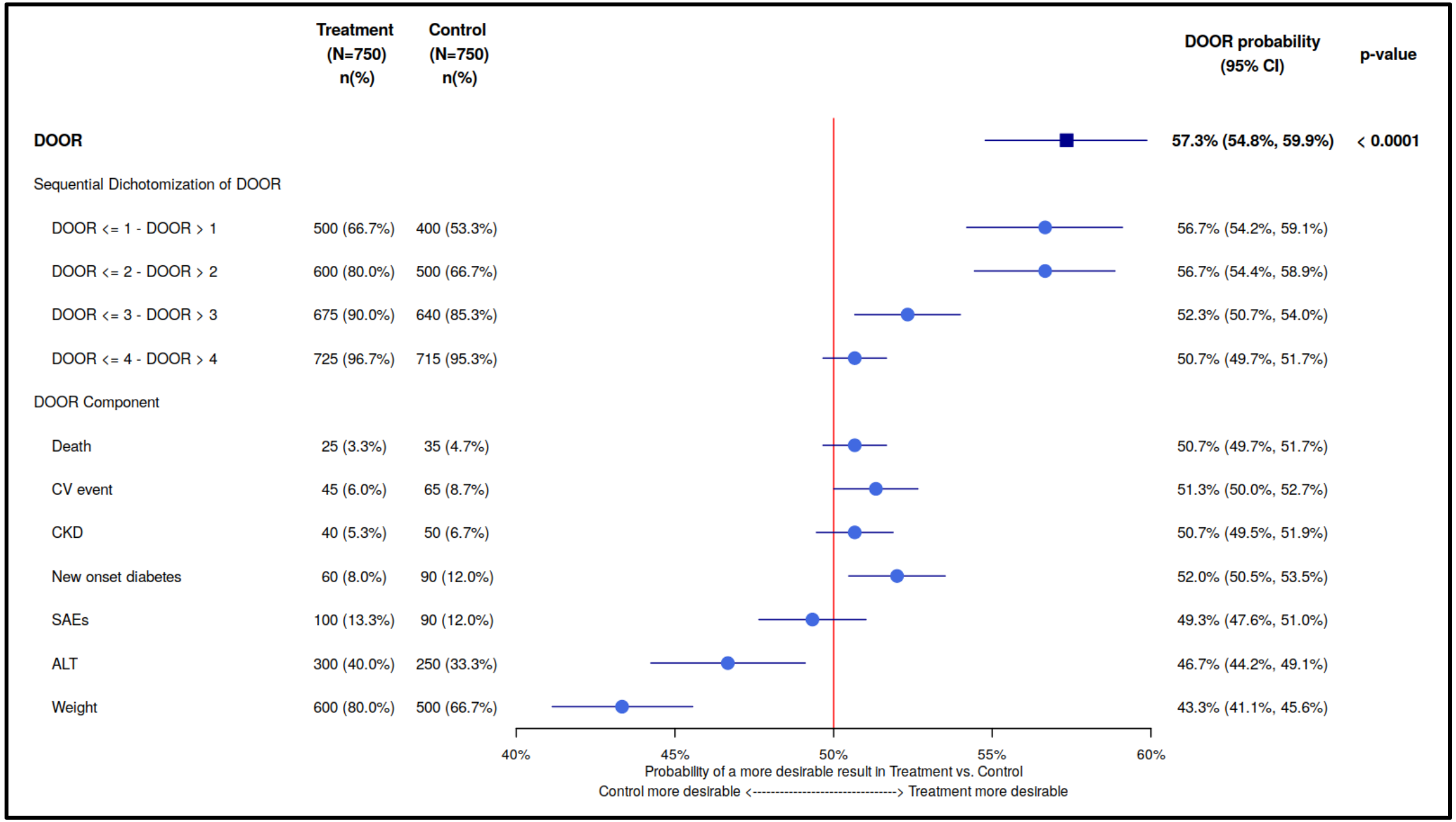
Question

- Suppose an intervention increases risk of death of 1 in 10 to 2 in 10
- RR=2. Very important.
- Suppose an intervention increases risk of death of 1 in 10,000 to 2 in 10,000
- RR=2. Nearly irrelevant.
- The confidence intervals for both cases are exactly the same.
- RR is a challenge to interpret
- Even trickier when interpreting multiple RRs and how they counter-balance benefits and harms

A DOORable NASH Trials?

- Important events
  - Death
  - CV event
  - CKD
  - New onset diabetes
  - Severe toxicities from therapy
  - ALT
  - Weight

	Treatment	
DOOR	**Treatment**	**Control**
1. Alive: 0 events		
2. Alive: 1 event		
3. Alive: 2 events		
4. Alive: >2 events		
5. Death		


Treatment (N=750) n(%)
Control (N=750) n(%)
DOOR probability (95% CI)
p-value
DOOR
57.3% (54.8%, 59.9%)
< 0.0001
Sequential Dichotomization of DOOR
DOOR <= 1 - DOOR > 1
500 (66.7%)
400 (53.3%)
56.7% (54.2%, 59.1%)
DOOR <= 2 - DOOR > 2
600 (80.0%)
500 (66.7%)
56.7% (54.4%, 58.9%)
DOOR <= 3 - DOOR > 3
675 (90.0%)
640 (85.3%)
52.3% (50.7%, 54.0%)
DOOR <= 4 - DOOR > 4
725 (96.7%)
715 (95.3%)
50.7% (49.7%, 51.7%)
DOOR Component
Death
25 (3.3%)
35 (4.7%)
50.7% (49.7%, 51.7%)
CV event
45 (6.0%)
65 (8.7%)
51.3% (50.0%, 52.7%)
CKD
40 (5.3%)
50 (6.7%)
50.7% (49.5%, 51.9%)
New onset diabetes
60 (8.0%)
90 (12.0%)
52.0% (50.5%, 53.5%)
SAEs
100 (13.3%)
90 (12.0%)
49.3% (47.6%, 51.0%)
ALT
300 (40.0%)
250 (33.3%)
46.7% (44.2%, 49.1%)
Weight
600 (80.0%)
500 (66.7%)
43.3% (41.1%, 45.6%)
40%
45%
50%
55%
60%
Probability of a more desirable result in Treatment vs. Control
Control more desirable <----------> Treatment more desirable

Freely-available Online Analysis Tool

https://methods.bsc.gwu.edu/

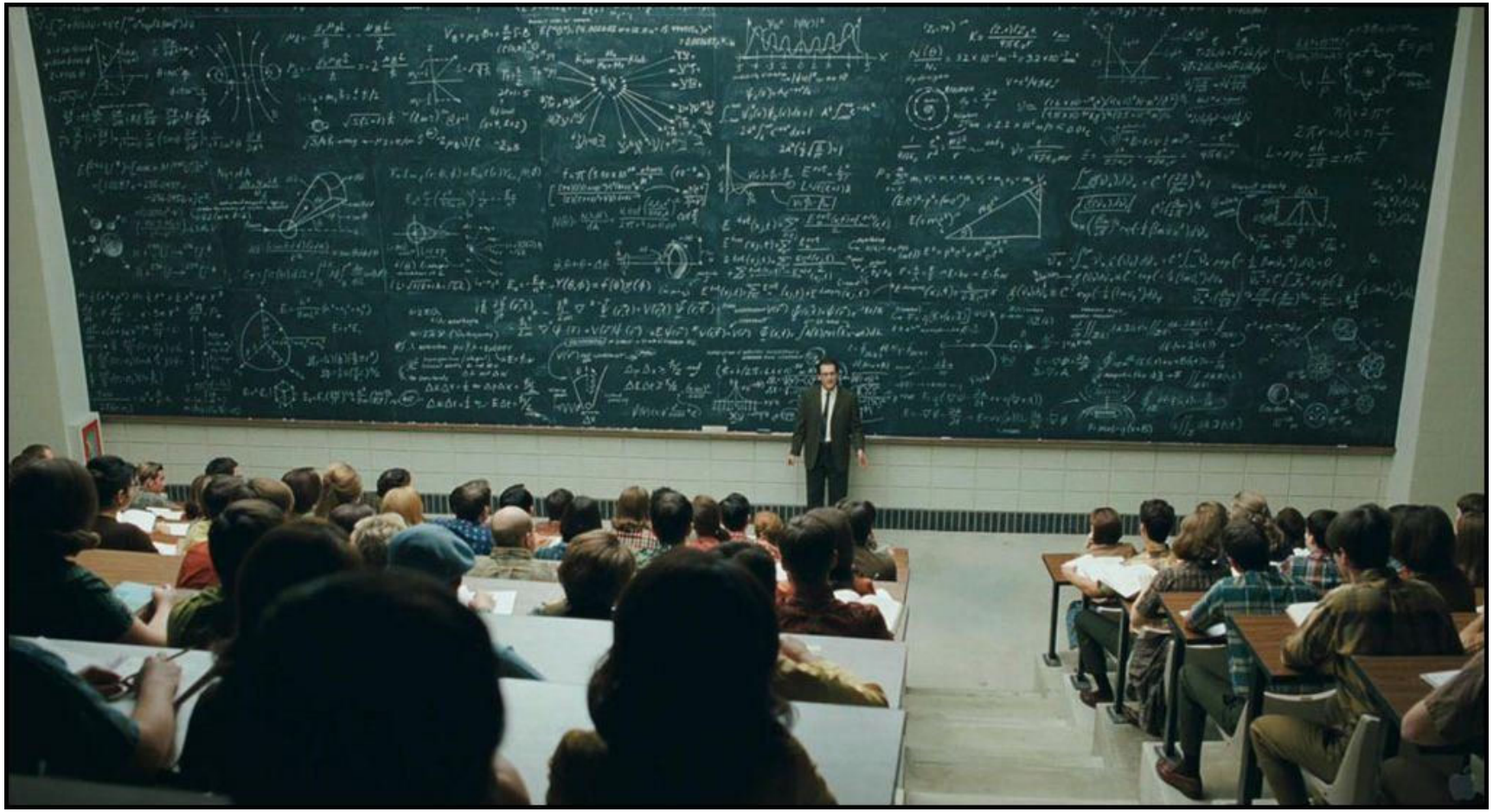
- Summary tables and graphics
- Expected comparative gain / loss in each category
- Partial credit

Conclusions

- The effects of interventions are multidimensional
- Use outcomes to analyze patients rather than patients to analyze outcomes
  - A closer reflection of the effects on patients
- DOOR
  - Effective tool for evaluating totality of patient-centric effects (benefit:risk)
  - May be tailored for NASH
  - Analysis of individual components is part of comprehensive DOOR analyses
  - May be sensitive due to recognition of finer gradations of patient response

Significant Contributors ($p<0.001$)

- Toshi Hamasaki
- Dean Follmann
- Dan Rubin
- Antibacterial Resistance Leadership Group
- ACTT-1 Investigators



I know that you will enthusiastically applaud now…

Because you are so relieved that it is over.

Thank you.