Risk Prediction for More Patient-Centered Evidence

David M. Kent, MD, MSc
Professor of Medicine, Neurology, Clinical and Translational Science,
Director, Predictive Analytics and Comparative Effectiveness (PACE) Center,
Institute for Clinical Research and Health Policy Studies, Tufts Medical Center

Personalized Medicine

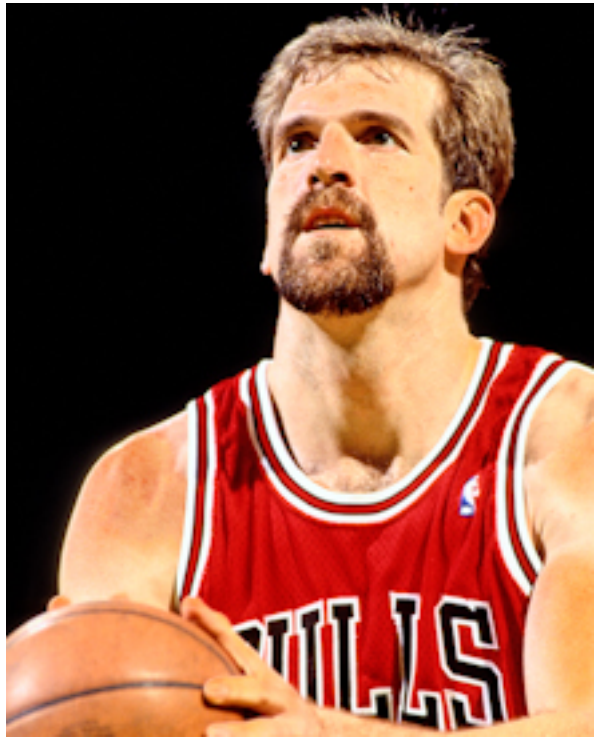
- “Personalized medicine is the practice of clinical decision-making such that the decisions made maximize the outcomes that the patient most cares about and minimizes those that the patient fears the most, *on the basis of as much knowledge about the individual’s state as is available.”*
  - **Pauker and Kassirer** N Engl J Med 316:250-258, **1987**

Evidence-based Medicine

- "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients."
  - Sackett JAMA 1996
- "Systematic review (with homogeneity) of RCTs" provide the best evidence
  - CEBM 2010

- What's best on average must be best for each individual.

The Fallacy of Division (Wennington's Fallacy)



Starters	PTS
B.J. Armstrong	16
Scottie Pippen	19
Michael Jordan	55
Toni Kukoc	3
Will Perdue	6
Reserves	**PTS**
Luc Longley	5
Corie Blount	2
Steve Kerr	5
Larry Krystkowiak	0
Bill Wennington	2
Pete Myers	0
Team Totals	**113**

Michael and I combined for 57 points"
-Bill Wennington, 1995

- It is potentially misleading to draw inferences about individuals based on aggregated characteristics of the (heterogeneous) group to which they belong.
- How do we estimate “individual” treatment effects?

Clinical Trial: "Box Score"

ACTUAL OUTCOME

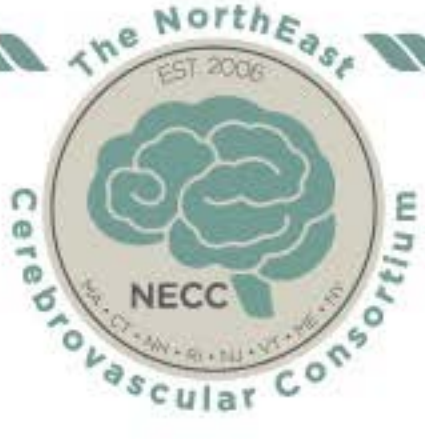
0 = alive
1 = dead

Subject Name	Without Treatment	With Treatment
SAM	0	
MARY		0
BOB	0	
BEN		0
CHRISTINE		0
NEIL	1	
MOHAMED		1
JENNIFER		1
PAUL	0	
NISHA	1	
MIGUEL	1	
LAYLA		0
PAUL	0	
EMANUEL		1
CHERYL		0
PATRICK	0	
OSCAR		1
JULIANNE	0	
THOMAS	0	
GEORGE		0

Individual Treatment Effects in a Deterministic Framework: Four possibilities

Without Treatment	With Treatment	
0	0	NO EFFECT
0	1	**HARM**
1	0	**BENEFIT**
1	1	NO EFFECT

Clinical Trial: "Box Score"



ACTUAL OUTCOME
COUNTER FACTUAL OUTCOME

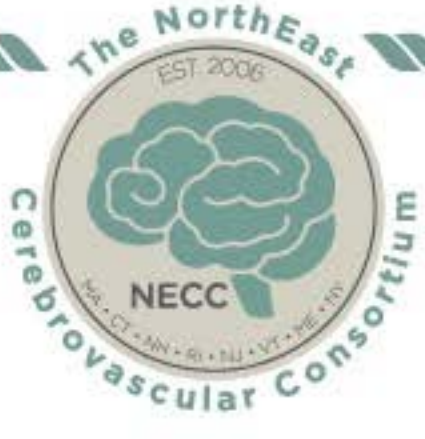
0 = alive
1 = dead

Subject Name	Without Treatment	With Treatment
SAM	0	1
MARY	0	0
BOB	0	0
BEN	1	0
CHRISTINE	1	0
NEIL	1	1
MOHAMED	1	1
JENNIFER	1	1
PAUL	0	1
NISHA	1	1
MIGUEL	1	1
LAYLA	1	0
PAUL	0	0
EMANUEL	1	1
CHERYL	0	0
PATRICK	0	0
OSCAR	1	1
JULIANNE	0	0
THOMAS	0	0
GEORGE	1	0

Individual Treatment Effects in a Deterministic Framework: Four possibilities

Without Treatment	With Treatment	
0	0	NO EFFECT
0	1	HARM
1	0	BENEFIT
1	1	NO EFFECT

Clinical Trial: "Box Score"



ACTUAL OUTCOME
COUNTER FACTUAL OUTCOME

0 = alive
1 = dead

Subject Name	Without Treatment	With Treatment	
SAM	0	1	← HARM
MARY	0	0	
BOB	0	0	
BEN	1	0	← BENEFIT
CHRISTINE	1	0	← BENEFIT
NEIL	1	1	
MOHAMED	1	1	
JENNIFER	1	1	
PAUL	0	1	← HARM
NISHA	1	1	
MIGUEL	1	1	
LAYLA	1	0	← BENEFIT
PAUL	0	0	
EMANUEL	1	1	
CHERYL	0	0	
PATRICK	0	0	
OSCAR	1	1	
JULIANNE	0	0	
THOMAS	0	0	
GEORGE	1	0	← BENEFIT

Clinical Trial: “Box Score”

Without Treatment	With Treatment
0	1
0	0
0	0
1	0
1	0
1	1
1	1
1	1
0	1
1	1
1	1
1	0
0	0
1	1
0	0
0	0
1	1
0	0
0	0
1	0

0 = alive
1 = dead

BENEFIT

Proportion Dead	**11/20 (55%)**	**9/20 (45%)**

Why do we fail to reliably detect HTE?

1. **Information failure**
   - Observable co-variates are totally unrelated to the causal determinants of HTE.
   - We just need better biomarkers, better genomics, better imaging
2. **Analytic failure**
   - Low power
   - Limitations of conventional (one-variable-at-a-time) subgroup analysis

Problems With Conventional Subgroup Analysis

- **Spurious False Positives**

“Positive” subgroup analyses subsequently shown to be false

Observation	Refutation
Aspirin is ineffective in secondary prevention of stroke in women[29,30]	31

Rothwell PM. *Lancet* 2005;365(9454):176-86.

Problems With Conventional Subgroup Analysis

- **Spurious False Positives**
- **Compare groups of patients that are more similar than disimilar.**
- **Individuals patients belong to many different subgroups.**

Interim Summary

- Determining the best treatment on average (the task of an RCT) is very different from determining the best treatment for an individual (the task of a good clinician) .
- Conventional subgroup analysis of clinical trials are typically inadequate and can also be misleading.

Why Risk Based Subgroup Analysis Should be Routine

David M. Kent, MD, MSc
Professor of Medicine, Neurology, Clinical and Translational Science,
Director, Predictive Analytics and Comparative Effectiveness (PACE) Center,
Institute for Clinical Research and Health Policy Studies, Tufts Medical Center

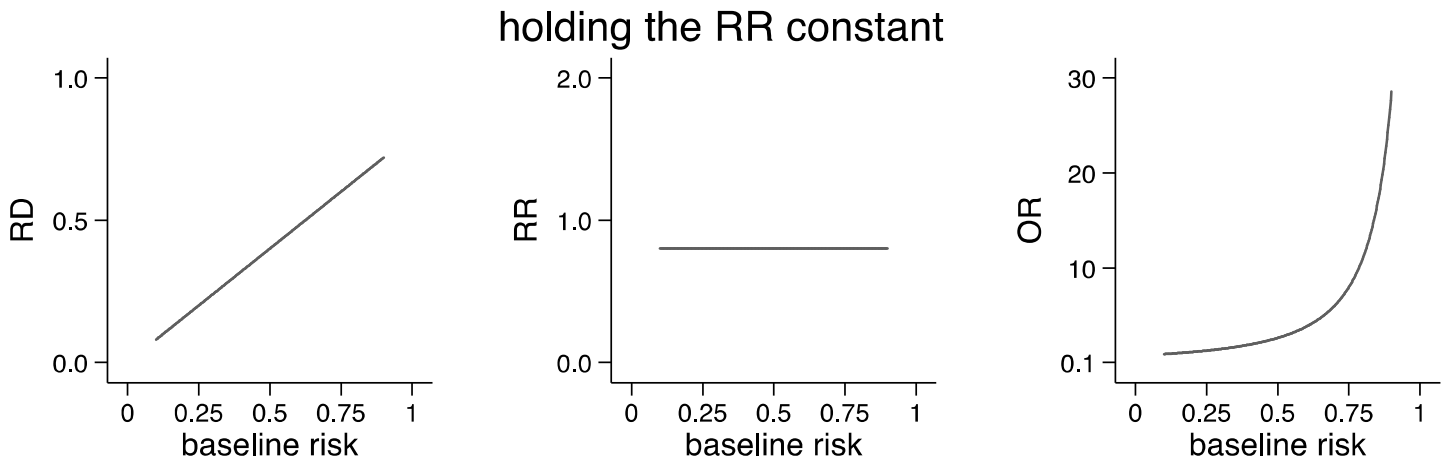
Why privilege risk-based HTE analysis?

- Risk is a known mathematical determinant of treatment effect.

Common Measures of Treatment Effect

Risk Reduction (RR)	Definition
Absolute RR	EER-CER
Relative RR	$1 - \frac{EER}{CER}$
Odds Ratio	$\frac{EER/(1-EER)}{CER/(1-CER)}$
CER=control event rate EER=experimental event rate	

An Illustration of Scale Dependence of HTE over Baseline Outcome Risk



Why privilege risk-based HTE analysis?

- Risk is a known mathematical determinant of treatment effect.
- When baseline risk heterogeneity is present (and the treatment effect is non-zero), there is always HTE.
- Risk provides a summary measure that takes into account multiple variables that are relevant; provides “patient-centered” evidence.

Figure 1: Distribution of Mortality Risk with Thrombolytic Thearpy in Patients with Acute Myocardial Infarction

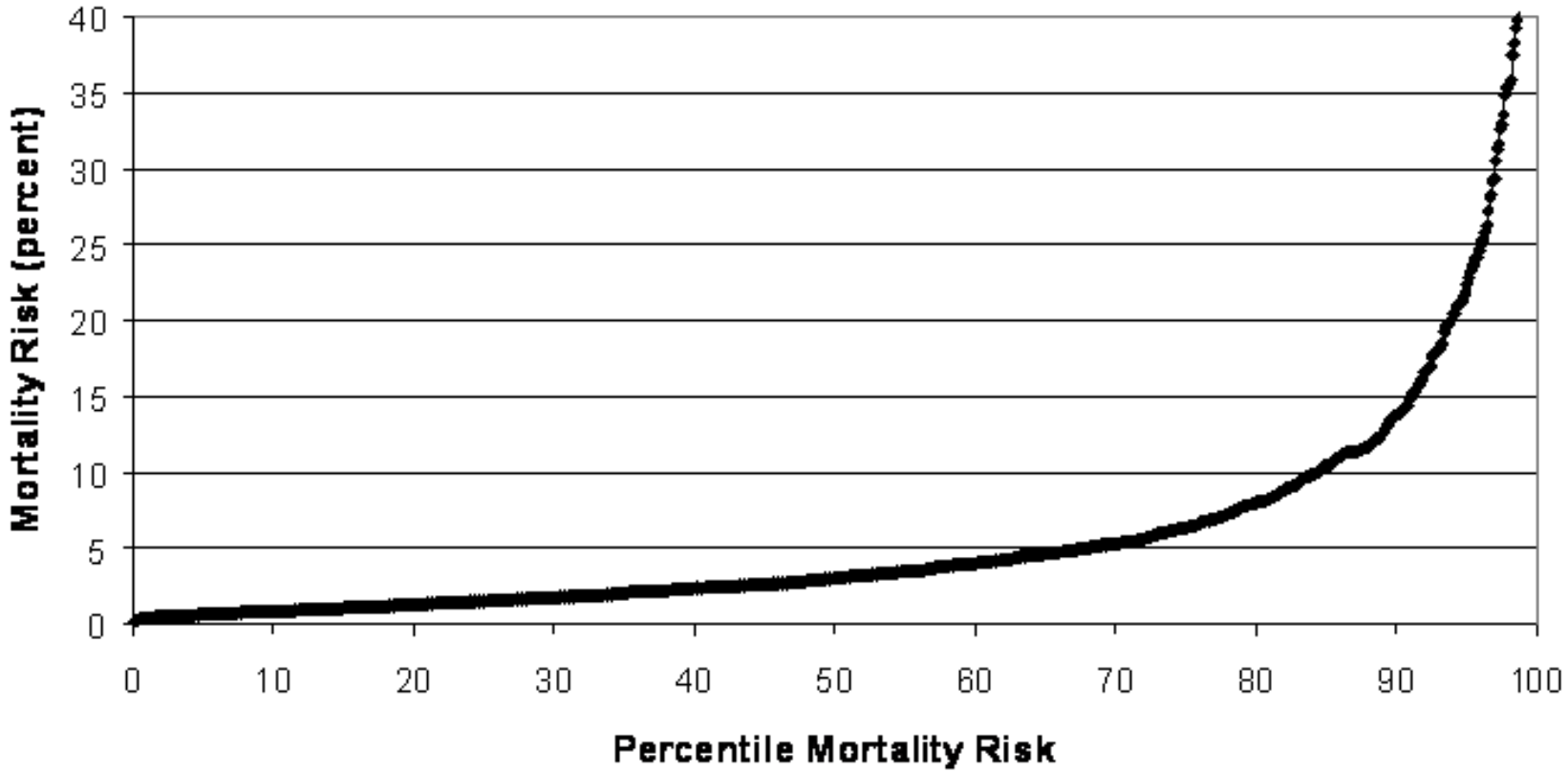


Figure 1: Distribution of Mortality Risk with Thrombolytic Thearpy in Patients with Acute Myocardial Infarction

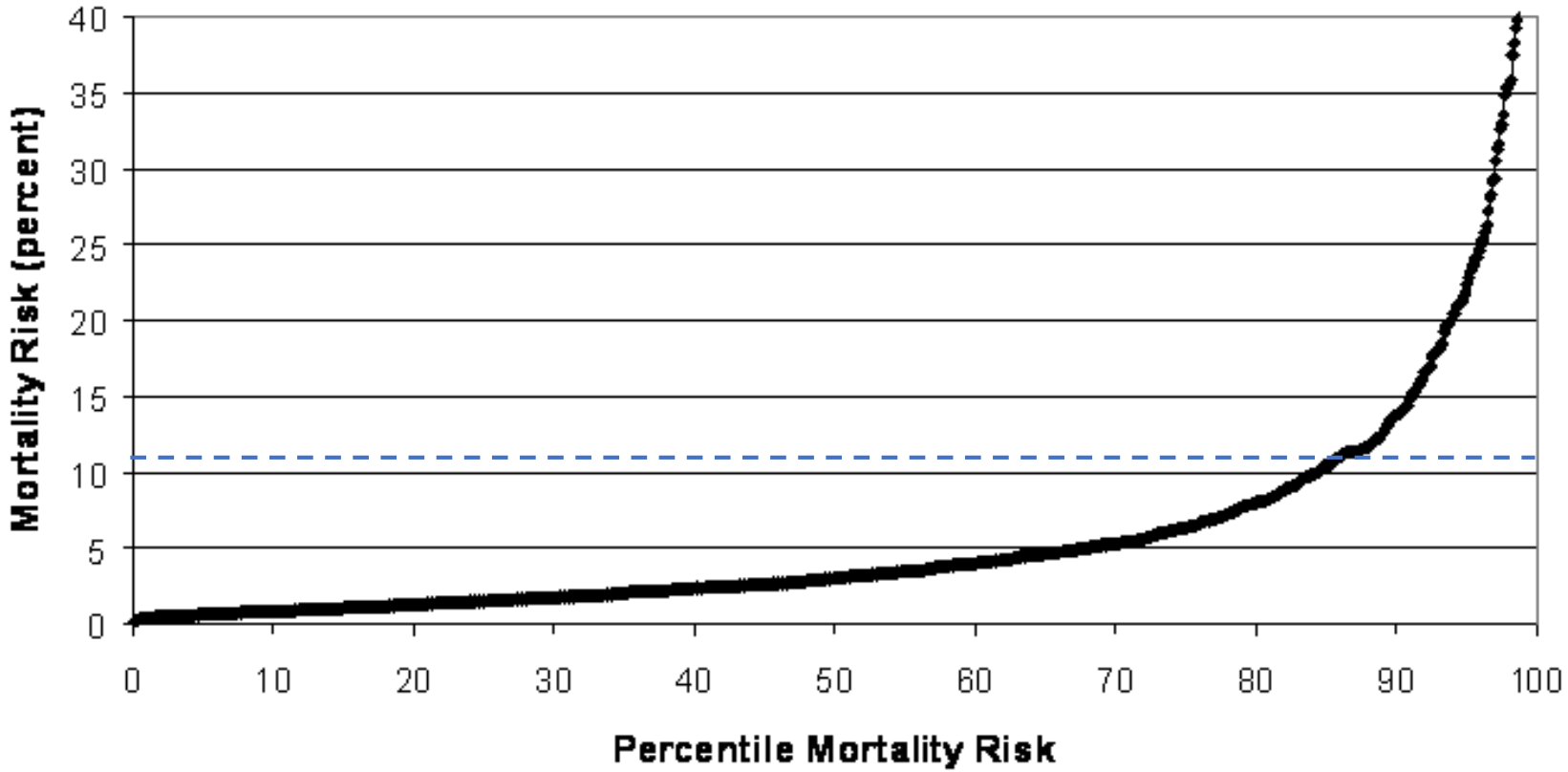
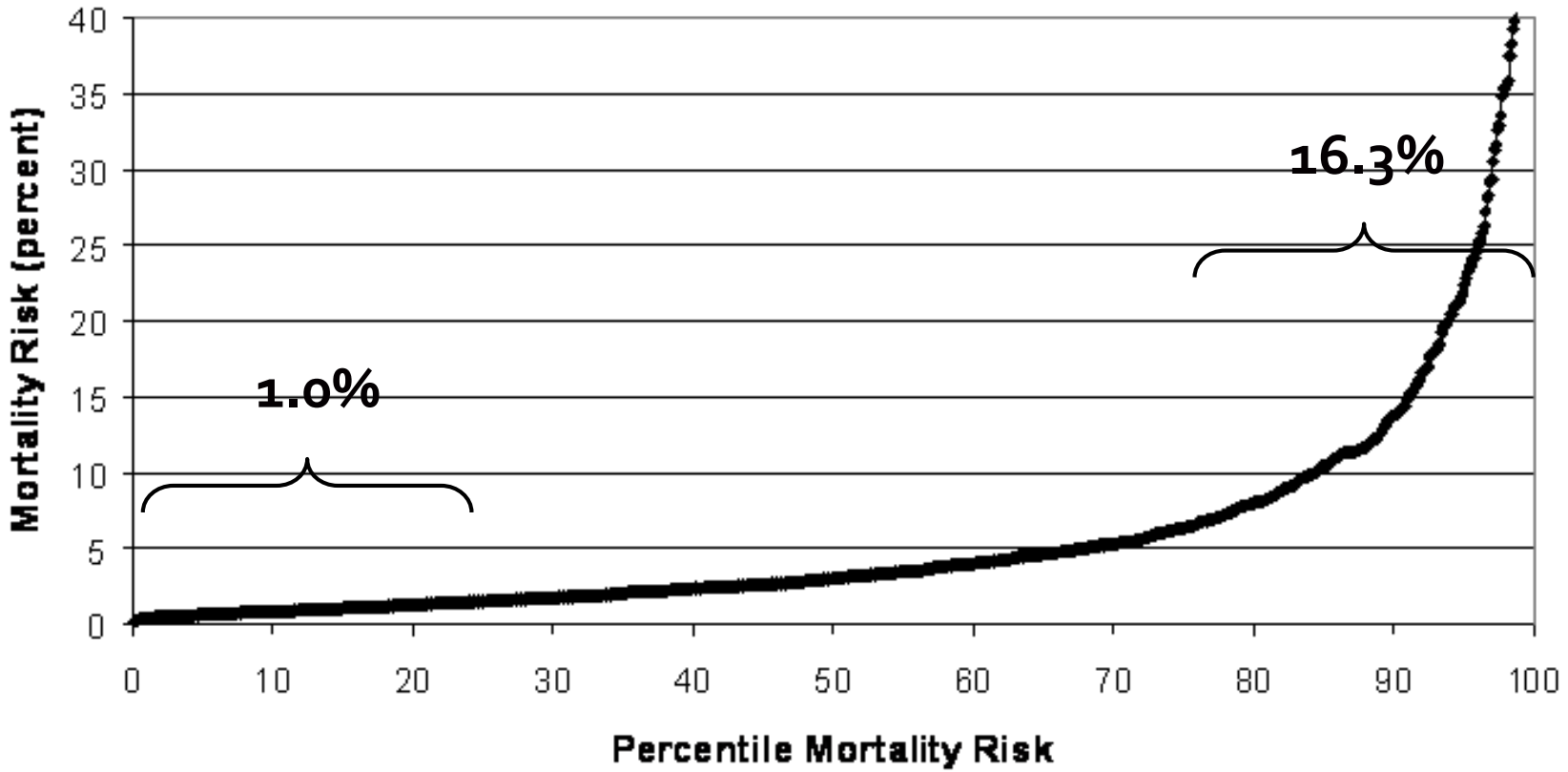
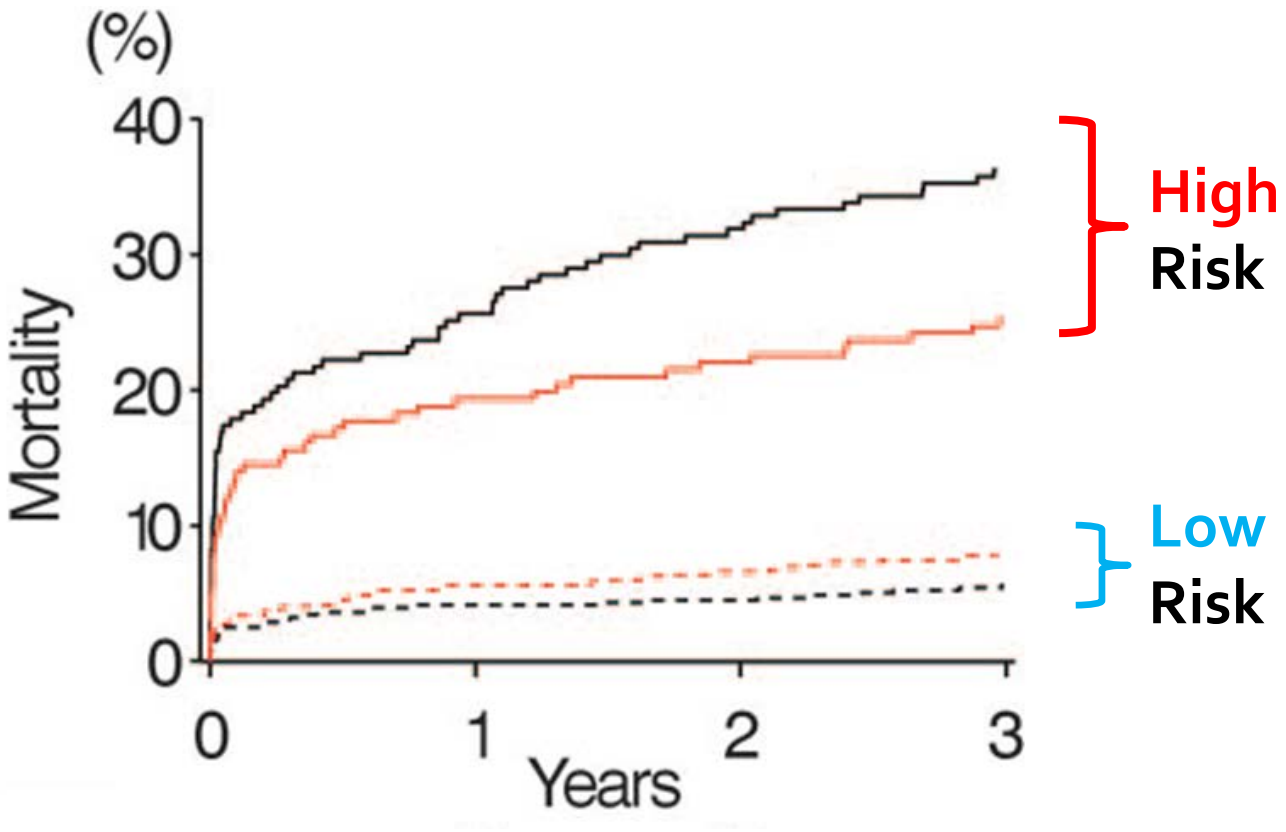


Figure 1: Distribution of Mortality Risk with Thrombolytic Thearpy in Patients with Acute Myocardial Infarction



DANAMI-2

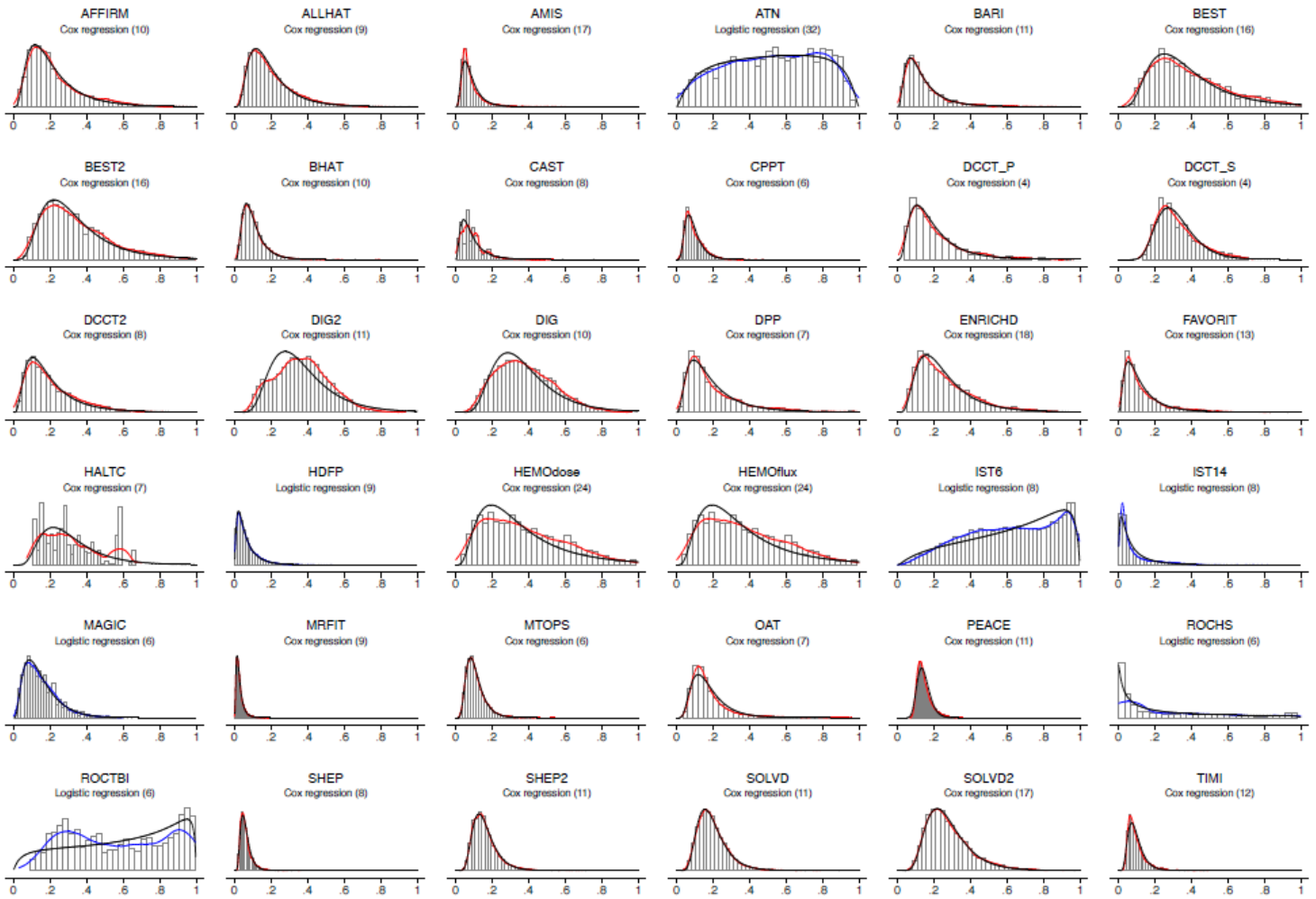


Number at risk

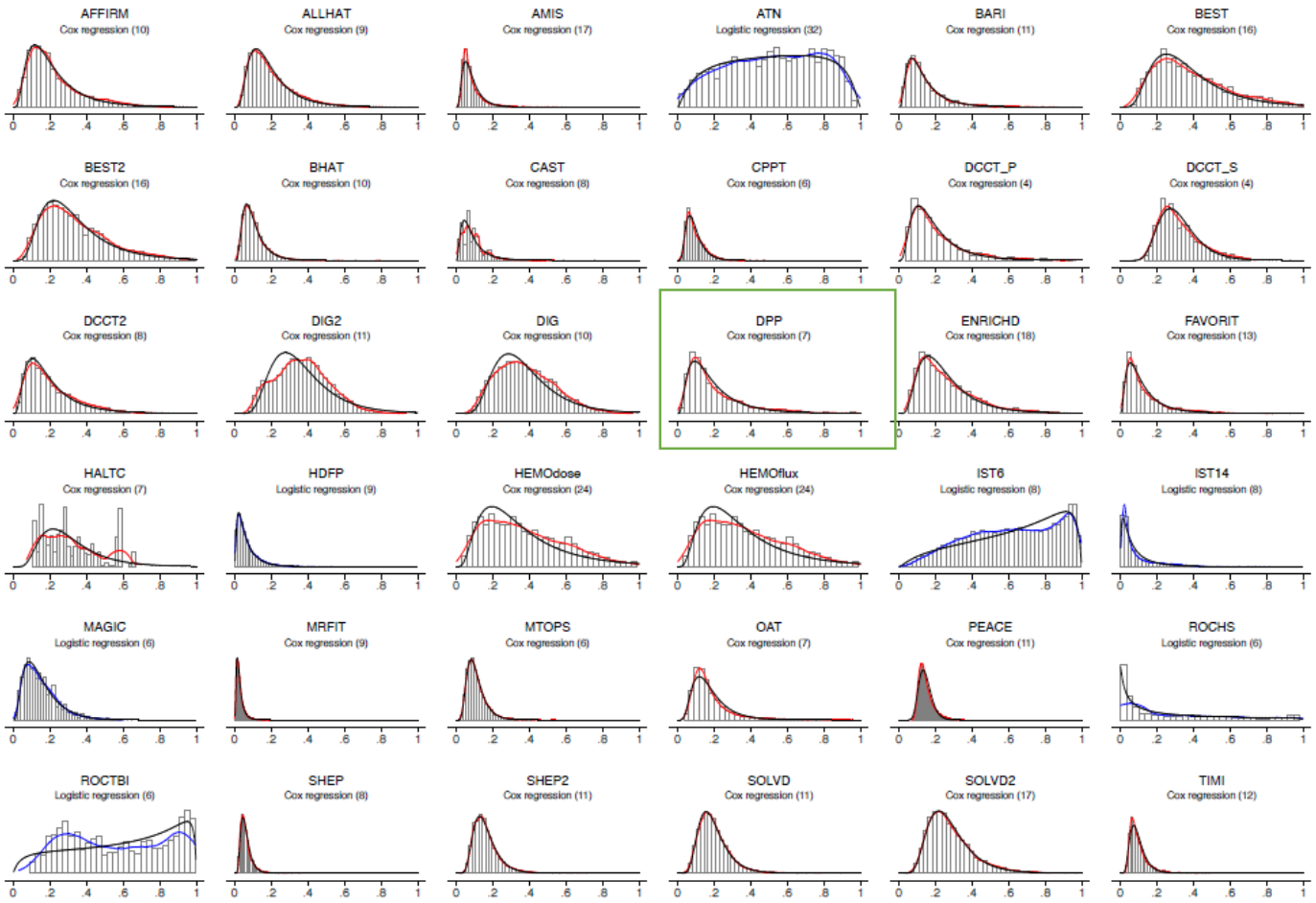
TIMI 0-4	Fx	556	533	531
	PA	578	546	540
TIMI ≥ 5	Fx	207	154	141
	PA	186	150	145

Thune JJ, et al. *Circulation* 2005,112:2017-2021.

Predicted Risk Distributions in RCTs



Predicted Risk Distributions in RCTs

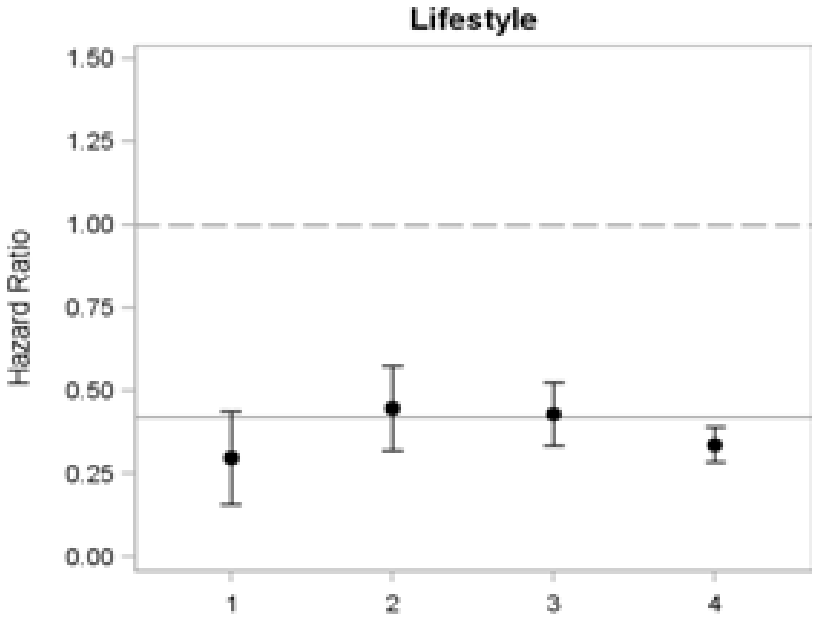


Diabetes Prevention Program (DPP) Randomized Controlled Trial

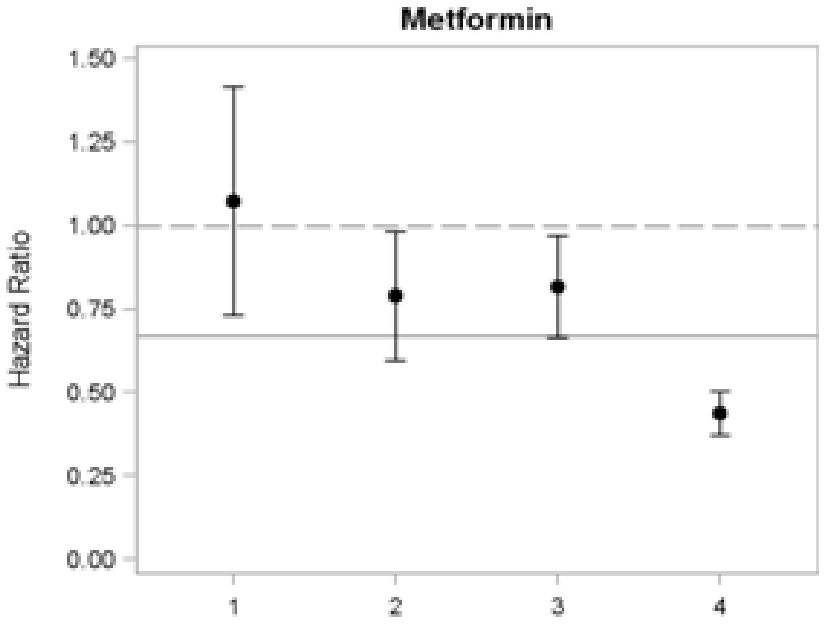
- Participants: 3060 nondiabetic persons with evidence of impaired glucose metabolism.
- Intervention: Intervention groups received metformin or a lifestyle-modification program.
- Main Outcome Measure: Development of diabetes

The DPP study was conducted by the DPP Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

DPP Risk Stratified Results

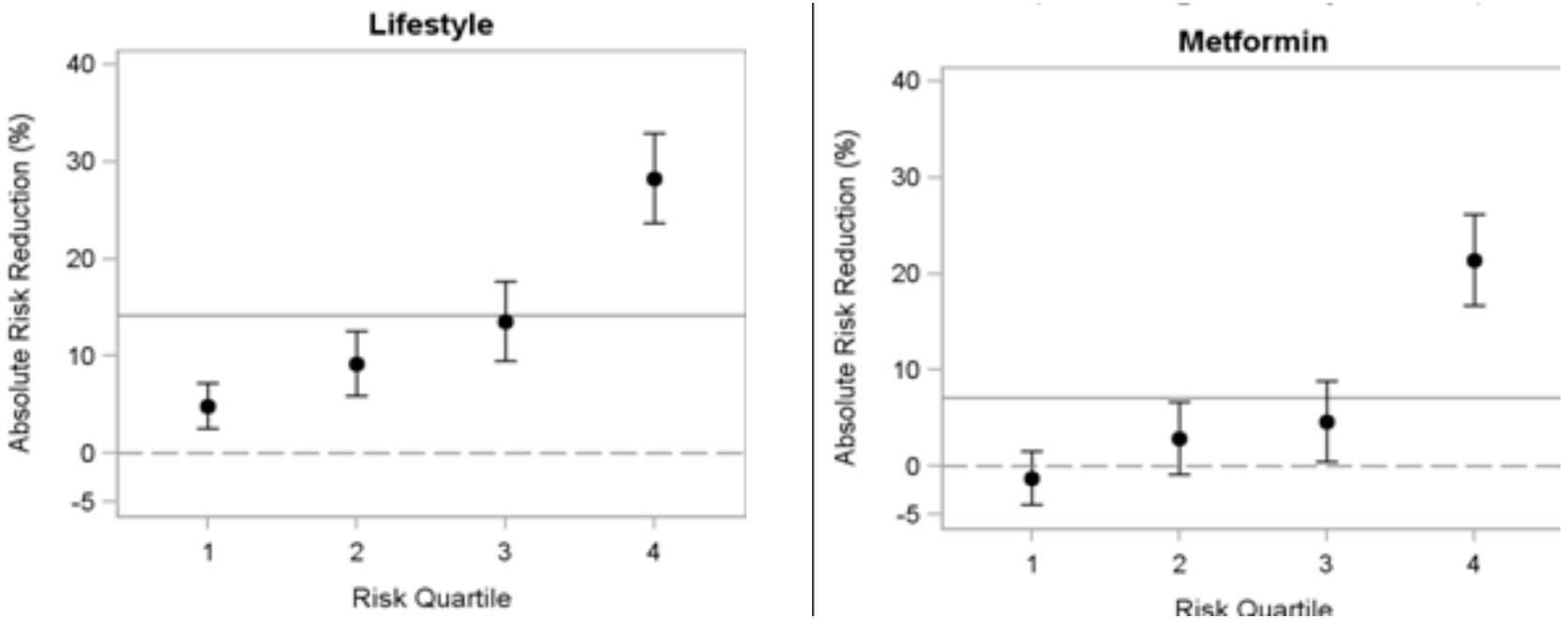


p value = NS

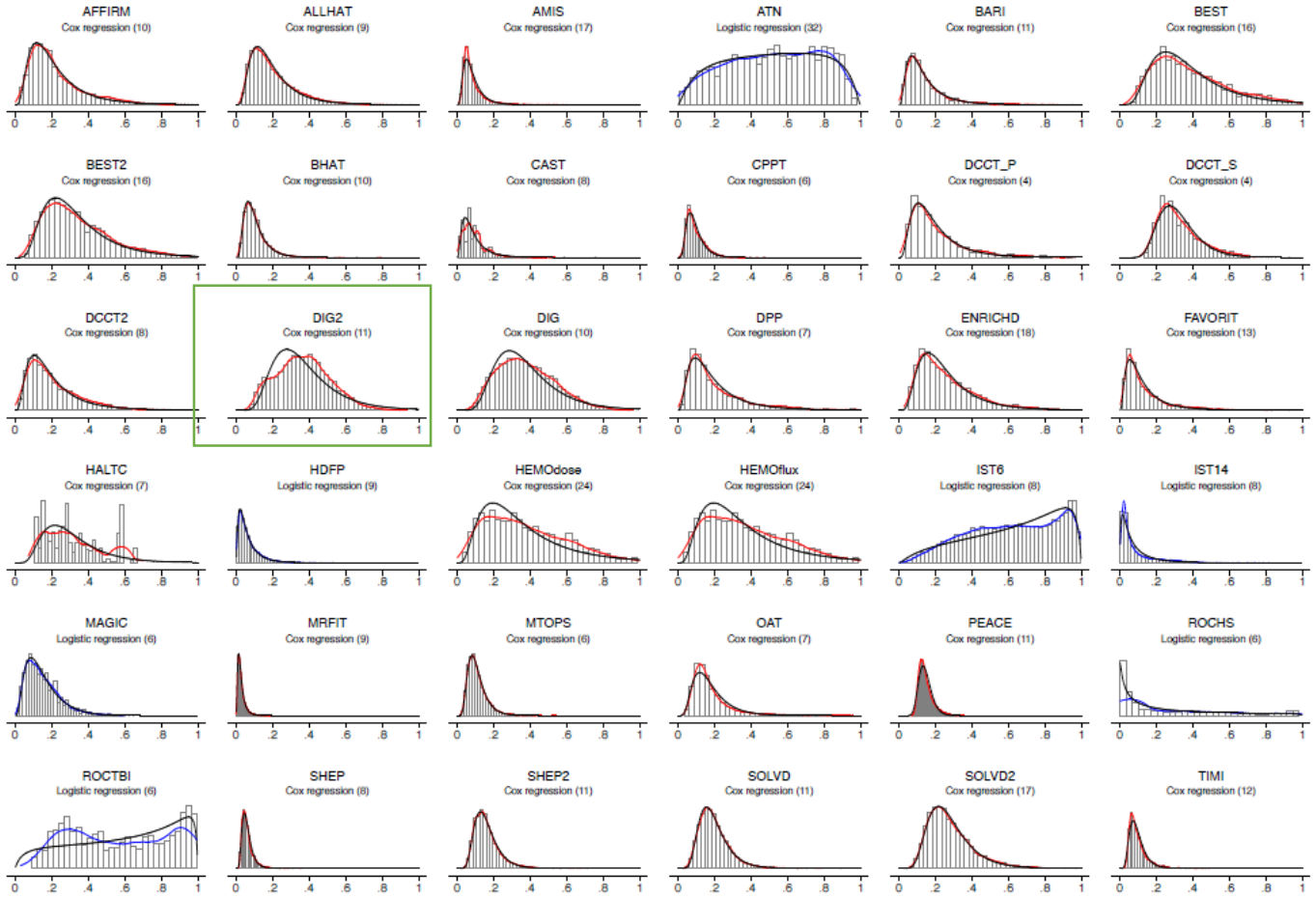


p value = 0.0008

DPP Risk Stratified Results



Predicted Risk Distributions in RCTs



Digitalis Investigator Group (DIG) Study

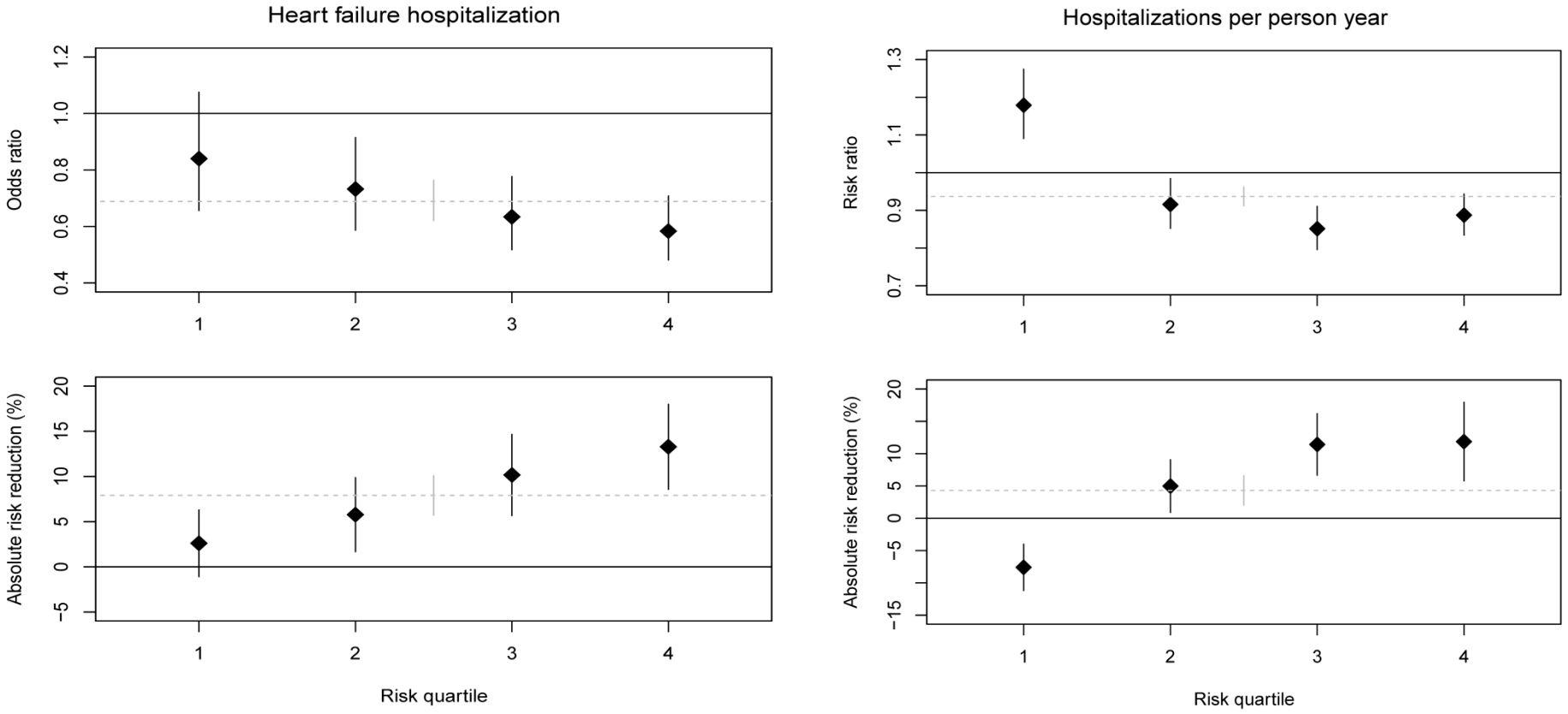
- Participants: Participants with HF and LVEF less than or equal to 45% (main DIG study, n=6800) or LVEF >45% (ancillary DIG study, n=988).

- Intervention: digoxin versus placebo

- Main Outcome Measure: Hospitalization due to worsening HF, all cause hospitalization

Upshaw JN et al. unpublished

DIG Risk Stratified Results



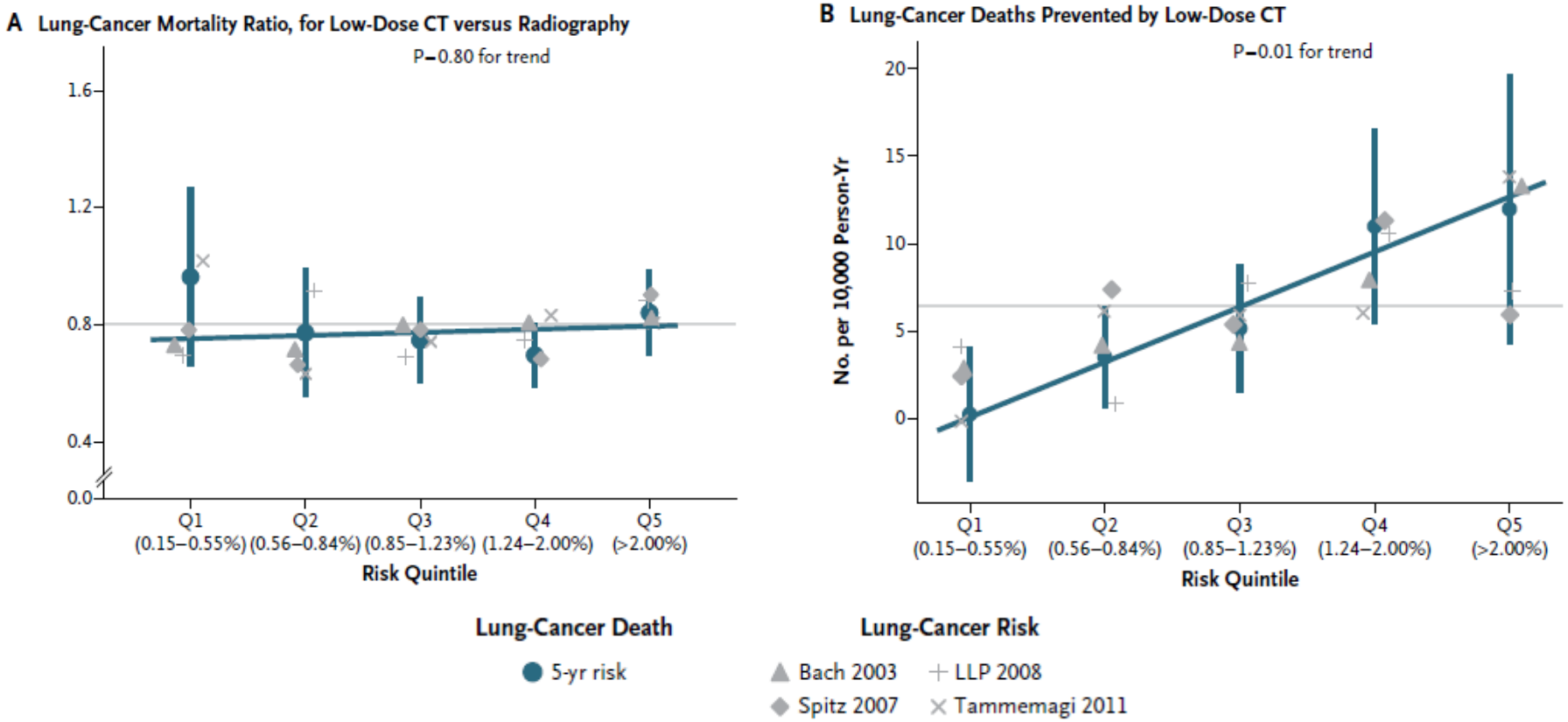
Upshaw JN et al. unpublished

National Lung Screening (NLST) Trial

- Participants: Smokers between the ages of 55 and 74 years with a minimum of 30 pack-years of smoking and no more than 15 years since quitting
- Intervention: Low-dose CT screening or chest radiography
- Main Outcome Measure: Lung-cancer deaths

Kovalchik SA et al. N Engl J Med 2013; 369: 245-54

NLST Risk Stratified Results



Kovalchik SA et al. N Engl J Med 2013; 369: 245-54

Risk based analyses can reveal counter-intuitive findings

- Overall effectiveness results may be driven by a relatively small group of influential (typically high risk) patients;
- The typical (median) risk patient is frequently at considerably lower risk than the overall average;
- The average benefit seen in the summary result often over estimates the benefit (on the RD scale) in most patients (and may obscure harm in many).

Clinical Conditions where Outcome Risk is Major Determinant of Clinically-Relevant HTE

CLINICAL CONDITION	INTERVENTION
Symptomatic carotid stenosis	Carotid endarterectomy
Non-valvular atrial fibrillation	Anticoagulation for primary prevention of stroke
Coronary artery disease	Coronary artery bypass grafting
Primary prevention of coronary artery disease	Blood pressure lowering Aspirin Lipid lowering
Acute coronary syndromes	Early invasive strategy (versus conservative) Clopidogrel (versus placebo) Enaxparin (versus unfractionated heparin)
ST-Elevation acute myocardial infarction	tPA (versus streptokinase) Percutaneous coronary intervention (versus thrombolytic therapy)
Severe sepsis	Drotrecogin alfa (activated protein C)
Pre-diabetes	Lifestyle intervention Metformin
Tobacco smoking	Lung cancer screening

Kent DM, et al. *Trials* 2010;11:85.

Summary

- Heterogeneity of outcome risk is ubiquitous.
- Heterogeneity of outcome risk inevitably gives rise to heterogeneity of treatment effect.
- One variable at a time subgroup analyses are inadequate (and prone to spurious false positive results).
- Risk based subgroup analyses can do better.

METHODOLOGY Open Access

Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal

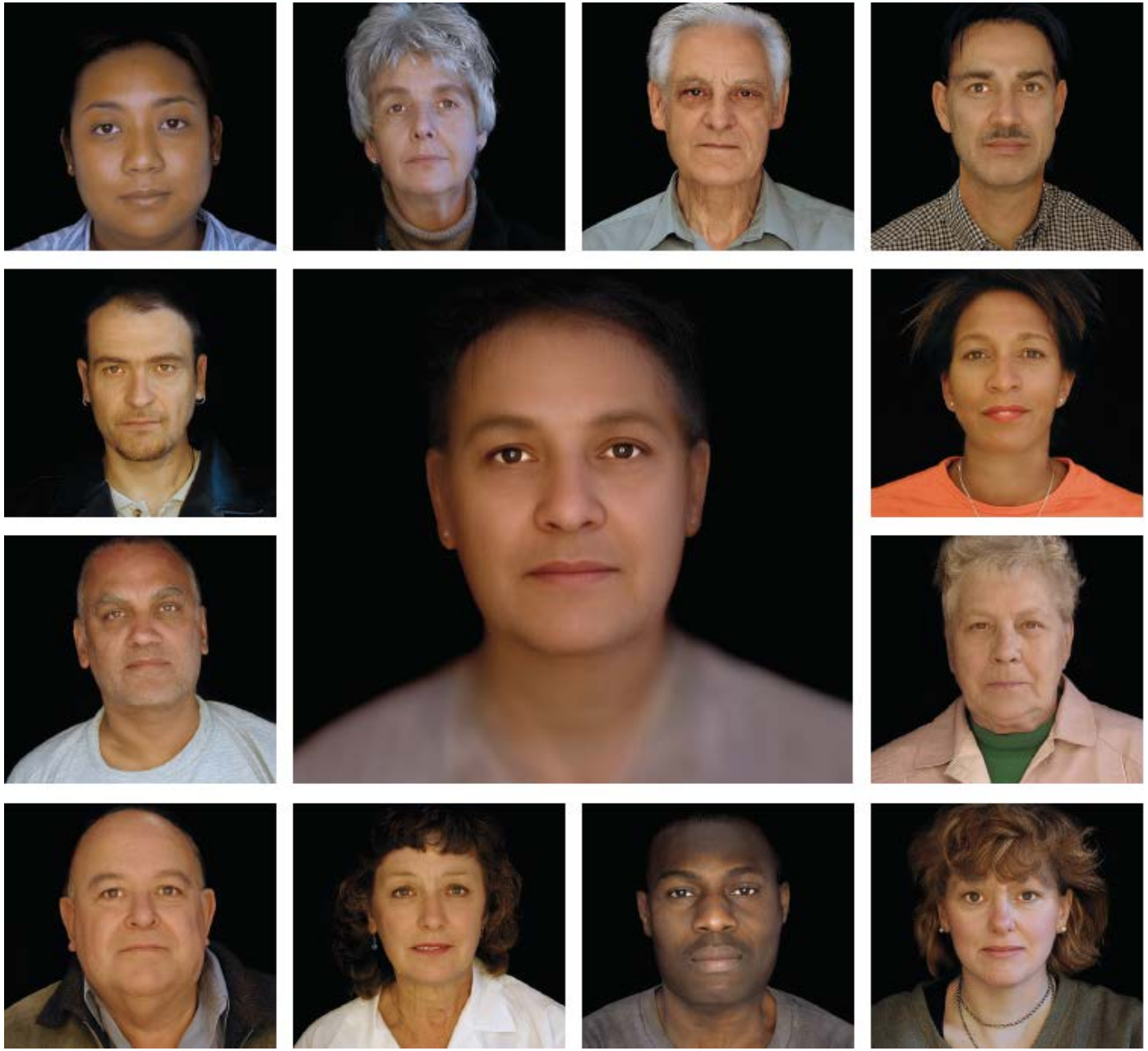
TRIALS

David M Kent[1*], Peter M Rothwell[2], John PA Ioannidis[1,3], Doug G Altman[4], Rodney A Hayward[5]

1. Evaluate and **report on the distribution of risk** in the overall study population and in the separate treatment arms of the study by using a risk prediction model or index.
2. Primary subgroup analyses should include reporting how relative and absolute risk reduction varies in a **risk-stratified analysis**.
3. Any additional **primary subgroup analysis should be pre-specified** and limited to patient attributes with strong a prior pathophysiological or empirical justification.
4. Conduct and **report on secondary (exploratory) subgroup analyses separate** from primary subgroup comparisons.
5. All analyses conducted must be reported and statistical testing of HTE should be done using **appropriate methods** (such as interaction terms) and avoiding over-interpretation.

Some (Really Important) Caveats

- Understanding how to model the relationship between risk and benefit, and how to estimate (absolute) individual treatment effects.
- How to capture the effects of important single-variable interactions, without including spurious interactions.
- How to include other dimensions (of risk and other things)



The NorthEast Cerebrovascular Consortium
EST 2006
NECC