# Decision Tools to Improve Personalized Care in Cardiovascular Disease: Moving the Art of Medicine towards Science

**Robert W. Yeh, MD, MSc** and **Daniel B. Kramer, MD, MPH**

Richard and Susan Smith Center for Outcomes Research in Cardiology, Division of Cardiology, Beth Israel Deaconess Medical Center and Harvard Medical School, both in Boston, MA

## Keywords

Personalized medicine; Decision Tool; Risk Score

> "If it were not for the great variability among individuals, medicine might as well be science, not an art."
>
> - Sir William Osler, 1892.

## Introduction

The proliferation of data collected in routine clinical care and advances in quantitative analytic methods have created immediate opportunities to individualize cardiovascular care through the use of *decision tools*. Such tools rely on the analysis of simple variables collected from medical records or direct patient observation to identify those with most to gain or lose from treatment, as determined by the collective experiences of similar patients treated previously. However, a number of obstacles limit the creation and utilization of decision tools. Here, we review the theoretical foundation of decision tools, highlight examples of successful efforts to model heterogeneity in treatment benefit, and suggest future goals to better utilize quantitative methods to personalize care.

## Identifying Heterogeneous Treatment Reponses

Arguments in favor of decision tools rest on three premises. First, differences in risk between patients must be identifiable by the tool more reliably than by clinical judgment alone (*identifiable heterogeneity*). Second, the identified risks should be modifiable by clinical decisions (*actionability*). Third, the tool should be able to be adopted into practice without disruption to patient care or work flow (*implementability*).

While the requirement for patient heterogeneity to enable personalized medicine may be self-evident, much of the literature that currently shapes cardiovascular practice fails to offer meaningful information to help clinicians identify or act on heterogeneity. Randomized trials typically report the overall treatment effect observed in the entire study sample as the top

Corresponding Author: Robert W. Yeh, 375 Longwood Ave, Suite 440, Boston, MA 02215. Phone: 617-667-8800, Fax: 617-632-7698, ryeh@bidmc.harvard.edu.

line result, giving an "average" result that may not represent the range of responses that could be observed in a diverse population. (Figure 1) Current examinations of heterogeneity within trials most often involve subgroup analysis based on individual stratification variables (e.g. sex), reported as a Forest plot examining the consistency of treatment effects across subgroups. Yet this approach commonly fails to identify meaningful treatment effect differences for a number of reasons. First, heterogeneity in treatment response may be best identified by stratification based on multiple factors rather than single variables.[1, 2] Next, studies are rarely powered to examine these effects, resulting in the inability to detect statistical interactions between subgroups. Finally, the identification of treatment effect heterogeneity within clinical trials, by convention, has generally examined interactions on the relative rather than absolute scale. Thus, subgroup analyses of clinical trials can fail to identify patient groups with substantially greater or lesser absolute treatment benefit than that observed in the overall study population.

## Dissociating Risk and Treatment Benefit

The need for decision tools also reflects important shortcomings in risk models, which have proliferated rapidly alongside the growth of registries and claims databases. The most valuable models accurately predict risk with high fidelity and clearly inform a clinical decision that can be made to mitigate that risk. Yet current risk models often fail to meet this standard, for several reasons. One, the events studied are frequently a mix of entities without a common causal pathway (e.g. all-cause readmission or generalized bleeding), which impairs the development of targeted interventions. Two, there may be no evidence that any intervention exists to mitigate the risk being predicted. For example, in-hospital mortality can be predicted with very high accuracy in patients presenting with acute myocardial infarction, but specific treatment strategies such as the use of hemodynamic support devices have not been consistently demonstrated to alter this risk. Finally, the use of risk scores to inform individual treatment decisions often use predicted risk (i.e. the likelihood of an event occurring) as a surrogate for the expected treatment effect (i.e. the change in the likelihood of the event occurring with vs. without treatment). This same fallacy may manifest in the opposite direction, with very high risk assumed to indicate futility. Risk scores are constructed in this manner often by necessity, as observational datasets may lack the ability to predict treatment benefit in an unbiased fashion. However, the conflation of patient risk with treatment benefit is often erroneous. The finding that the $CHA_2DS_2$-VASC score predicts stroke even in patients *without* atrial fibrillation[3] highlights the potential pitfalls of equating risk and treatment benefit, as (for example) left atrial appendage exclusion would likely not improve the outcomes of patients at high risk for stroke from multiple potential etiologies

## Towards Modeling Treatment Effect Rather Than Risks

The strengths and shortcomings of clinical trials and risk models have informed the development of decision tools currently in practice, several of which can identify heterogeneity, separate risk and benefits, and provide actionable information. For example, the DAPT Score guides decisions regarding the optimal duration of dual antiplatelet therapy.[4] This score was developed from randomized trial data, used multiple variables to

define the subgroups with the greatest or least benefit from treatment, and measured treatment differences on the absolute rather than relative scale. This approach integrated the combined risks of bleeding and ischemic events in an attempt to uncouple these correlated risks, providing information to evaluate potential benefit and harm of therapy, as opposed to merely the level of risk of the patient. The analysis of randomized trial evidence in such a fashion represents a break from usual conventions, and, we believe, offers the best opportunity for trial evidence to support personalized treatment decisions.

For decision tools to influence practice, external validation is critical. The DAPT Score was validated using non-randomized data, and its ability to predict events was modest when measured by the statistical metrics such as the area under the receiver operating characteristic curve (or C-statistic). However, randomized studies demonstrating that the use of decision tools has led to an improvement in clinical outcomes compared with usual care would be the true test of their value, but have not been performed.

## Fulfilling the Promise of Personalized Decision Tools

The challenges in identifying heterogeneity, separating risks and benefits, and demonstrating improved outcomes represent substantial hurdles for developing new decision tools. Yet to truly individualize care, personalized decision tools should enhance not only clinician understanding, but also help to inform patients and promote engagement in shared decision making. While decision aids may support these goals, creating instruments that integrate more quantitative models while remaining acceptable to practitioners remains difficult – but achievable.[5] Identifying genetic risk may play an important role in future applications of personalized medicine, but more achievable goals are currently within reach. Moving beyond conventions of clinical trial subgroup analysis, developing tools that quantify an individual's expected treatment effect rather than risk alone, and implementing these in a manner that can engage physician-patient dialogue may ultimately lead to tangible benefits to individuals while collectively improving population health.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## References

1. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC medical research methodology. 2006; 6:18. [PubMed: 16613605]

2. Hayward RA, Kent DM, Vijan S, Hofer TP. Reporting clinical trial results to inform providers, payers, and consumers. Health affairs (Project Hope). 2005; 24:1571–81. [PubMed: 16284031]

3. Mitchell LB, Southern DA, Galbraith D, Ghali WA, Knudtson M, Wilton SB. Approach investigators. Prediction of stroke or TIA in patients without atrial fibrillation using CHADS2 and CHA2DS2-VASc scores. Heart. 2014; 100:1524–30. [PubMed: 24860007]

4. Yeh RW, Secemsky EA, Kereiakes DJ, Normand SL, Gershlick AH, Cohen DJ, Spertus JA, Steg PG, Cutlip DE, Rinaldi MJ, Camenzind E, Wijns W, Apruzzese PK, Song Y, Massaro JM, Mauri L. Development and Validation of a Prediction Rule for Benefit and Harm of Dual Antiplatelet Therapy Beyond 1 Year After Percutaneous Coronary Intervention. Jama. 2016; 315:1735–49. [PubMed: 27022822]

5. Spertus JA, Decker C, Gialde E, Jones PG, McNulty EJ, Bach R, Chhatriwalla AK. Precision medicine to improve use of bleeding avoidance strategies and reduce bleeding in patients undergoing percutaneous coronary intervention: prospective cohort study before and after implementation of personalized bleeding risks. Bmj. 2015; 350:h1302. [PubMed: 25805158]
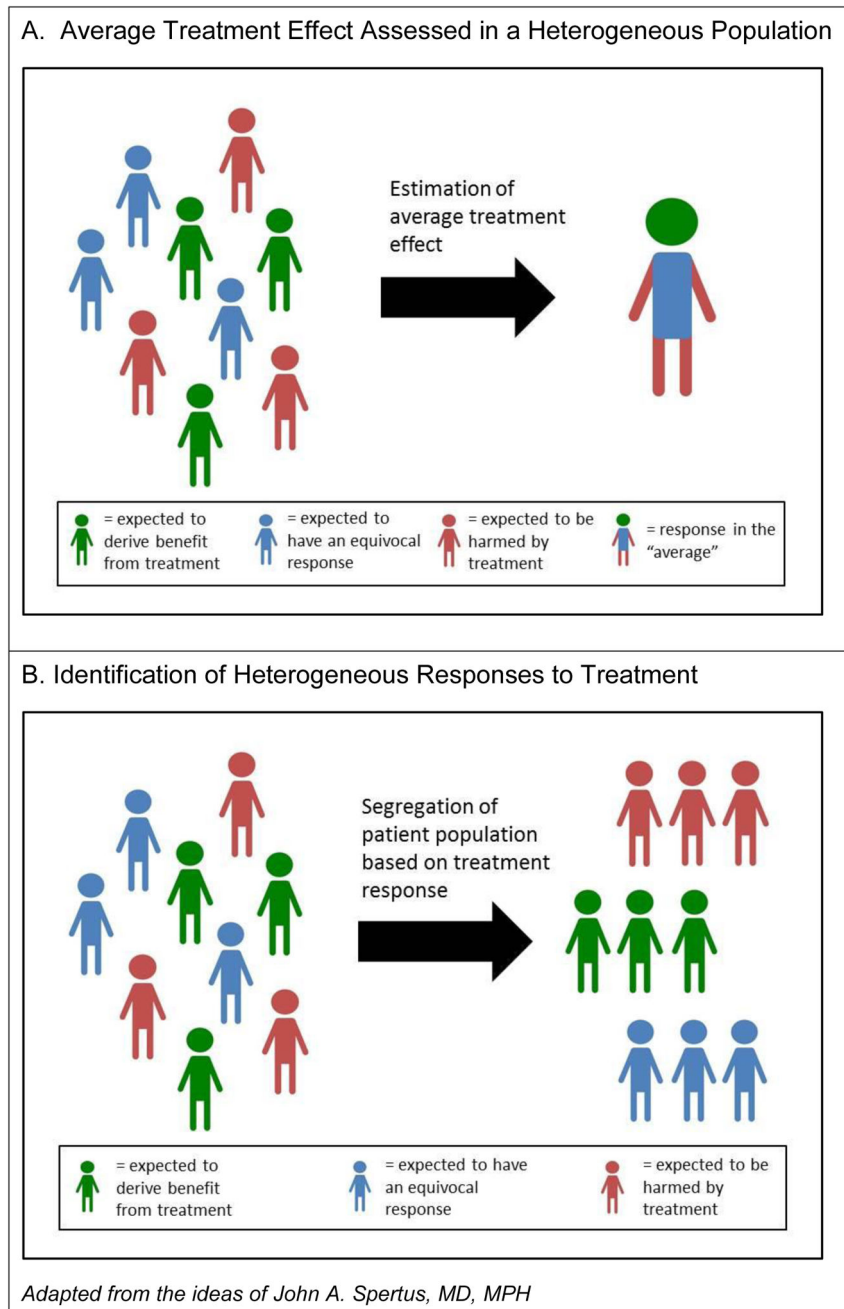
**Figure 1.**

**Figure 1A.** Conventional presentation of randomized clinical trial results focuses on average treatment effects, despite there being a potentially wide range of treatment responses among patients. The resulting estimate may not accurately describe the expected response for individual patients.

**Figure 1B.** Analysis of data that allows for the identification of subgroups of patients with the most to gain or lose from treatment can serve as the basis for new personalized decision tools.