



Published in final edited form as:

*J Am Coll Cardiol.* 2010 August 3; 56(6): 490–498. doi:10.1016/j.jacc.2009.11.102.

## Surgical Treatment for Ischemic Heart Failure (STICH) Trial Enrollment

**Robert H. Jones, MD, Harvey White, MB, ChB, DSc, Eric J. Velazquez, MD, Linda K. Shaw, Ricardo Pietrobon, MD, PhD, Julio A. Panza, MD, Robert O. Bonow, MD, George Sopko, MD, Christopher M. O'Connor, MD, and Jean-Lucien Rouleau, MD**

Division of Cardiothoracic Surgery/Department of Surgery, Duke University Medical Center, Duke Clinical Research Institute, Durham, North Carolina (RJ) Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand (HW) Division of Cardiovascular Medicine/Department of Medicine, Duke University Medical Center, Duke Clinical Research Institute, Durham, North Carolina (EV, CO) Duke Clinical Research Institute, Durham, North Carolina (LS) Department of Surgery, Duke University Medical Center; Department of Health Services Research, Duke-NUS (RP) Department of Medicine-Cardiovascular Disease, Washington Hospital Center, Washington, D.C. (JP) Division of Cardiology, Northwestern University Feinberg School of Medicine, Northwestern Memorial Hospital, Chicago, Illinois (RB) Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland (GS) Institut de Cardiologie de Montreal, University of Montreal, Montreal, Canada (JR)

### Abstract

**Objective**—To assess the influence of enrolling site location and enrollment performance on generalizability of STICH trial results.

**Background**—The international Surgical Treatment for Ischemic Heart Failure (STICH) trial seeks to define the role of cardiac surgery for ischemic cardiomyopathy patients.

**Methods**—Baseline characteristics of 2,136 randomized STICH patients were entered into a multivariable equation created using the Duke Databank for Cardiovascular Diseases to predict their five-year risk of death without cardiac surgery. Patients ordered by increasing predicted risk were assigned to 1 of 32 risk at randomization (RAR) groups created to share 1/32 of total predicted deaths. Numbers of patients sharing the same RAR group were compared between higher- and lower-enrolling site groupings and for countries tending to enroll high-risk or low-risk patients.

**Results**—Country of enrollment was a stronger determinant of risk diversity than site enrollment performance among patients enrolled at 127 sites in 26 countries. Mean RAR differences among countries ranged from 9.4 (Singapore) to 18.6 (Germany). However, 1,614 (76%) of 2,136 patients from countries enrolling lower-risk patients shared the same RAR group with patients from countries enrolling higher-risk patients. Baseline characteristics responsible for risk differences of

---

© 2011 American College of Cardiology Foundation. Published by Elsevier Inc. All rights reserved.

Address for Correspondence: Robert H. Jones, M.D. Mary and Deryl Hart Professor of Surgery P. O. Box 2986, Duke University Medical Center Durham, NC 27710 Telephone: 919-668-8357 Fax: 919-684-5700 jones060@mc.duke.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors have no relationships with industry or financial conflicts of interest to disclose.

patients enrolled in the two country groupings were sufficiently similar to exert little influence on clinical decision-making.

**Conclusions**—STICH randomized patients are characterized by a continuous spectrum of risk without discordant dominance from any site or country. Clinical site diversity promises to enhance generalization of STICH trial results to a broad population of ischemic cardiomyopathy patients.

### Keywords

coronary artery bypass grafting; surgical ventricular reconstruction; randomized clinical trial

## INTRODUCTION

The National Institutes of Health funded international Surgical Treatment for Ischemic Heart Failure (STICH) trial addresses two hypotheses to assess the value of cardiac surgery in patients with ischemic cardiomyopathy (1). The recently reported outcome of the surgical ventricular reconstruction (SVR) hypothesis confirmed that adding SVR to coronary artery bypass grafting (CABG) in patients with anterior left ventricular (LV) dysfunction decreased LV volume more than CABG alone but did not reduce death or cardiac hospitalization (2). Cardiovascular specialists have raised questions about whether the spectrum of severity of disease at time of randomization was sufficiently broad to include all patients who might receive survival benefit from SVR (3). Data from patients now being followed in the surgical revascularization hypothesis will definitively answer whether or not CABG added to evidence-based medical therapy (MED) increases survival of ischemic cardiomyopathy patients. This answer will clarify whether an ischemia diagnosis and interventional treatment should be pursued in patients presenting with heart failure symptoms or asymptomatic systolic dysfunction. The importance of STICH trial findings to guide appropriate care decisions for future patients demands an answer to the question – How generalizable are STICH trial results to the patients I see?

STICH protocol enrollment criteria ensured patient safety and excluded patients without ischemic cardiomyopathy as the most likely cause of death during follow-up. Thereafter, responsible physicians selected one of three randomization strata, each of which included possible randomization to a cardiac surgical option with significant short-term risk. Patient enrollment was difficult. Variation in equipoise among clinicians and clinical site diversity of structure and care processes produced a broad range of randomization rates. Incremental geographic expansion over three years caused enrollment duration to vary among sites. Clinical site location has been shown to influence characteristics of enrolled populations (4). That STICH was among the top 10 (7.9%) of 129 ongoing clinical trials during 2005 randomizing patients in 20 or more countries (5) emphasizes the need to assess whether enrolling site diversity helped or hindered generalization of STICH conclusions. Diversity of STICH sites could have aggregated multiple subgroups into discontinuous patient groupings with dissimilar risk. Alternatively, random individual patient diversity both within and among sites could produce a population with a smooth gradation of risk severity. A patient-based metric of baseline risk is needed to document the influence of enrollment decisions made by many physicians in different clinical care environments.

This report describes the STICH risk at randomization (RAR) index that reflects baseline clinical risk of each patient in context with the full risk spectrum of all randomized patients. Clinical site enrollment rates and geographic location influence on the total randomized patient cohort risk spectrum will be compared using the numbers of randomized patients who share the same RAR index number when grouped by clinical site characteristics.

Clinical data tabulated from RAR groups created to show the greatest baseline risk discordance will place the variability among enrolling sites in clinical context.

## METHODS

### Clinical Site and Patient Recruitment

In 2002, 23 United States, 7 Canadian, and 2 European clinical sites selected for expertise in clinical research and outstanding care of patients with ischemic cardiomyopathy were activated. Persistently low patient enrollment required incremental expansion that concluded in 2005 with 171 activated clinical sites. Clinical sites were expected to enroll patients into both hypotheses except for the last sites invited in Argentina, India, and the United Kingdom. These sites were recruited to enroll patients only into the surgical revascularization hypothesis because their help was not needed to attain the 1000-patient SVR hypothesis recruitment goal. During the three years of incremental expansion, 44 activated clinical sites were deactivated by mutual consent because of their inability to enroll any patient.

LV ejection fraction of  $\leq .35$  and coronary artery disease (CAD) amenable to CABG were entry criteria for all STICH patients. Major cardiac disease exclusions were a recent myocardial infarction and need for aortic valve replacement. Any non-cardiac disease sufficiently severe to dominate expected three-year longevity excluded patients. Study design required all STICH patients to receive MED. Patients with Canadian Class III or greater ongoing angina on MED and/or  $\geq 50\%$  left main stenosis were not eligible for randomization to MED as their only treatment. Eligibility for SVR required dominant anterior akinesia or dyskinesia amenable to SVR. Patients ineligible for SVR but eligible to receive MED with or without CABG were randomized in Stratum A. Patients ineligible for MED alone but eligible to receive CABG with or without SVR were randomized in Stratum C. Patients eligible for all three treatments were randomized in Stratum B. All Stratum A and Stratum B patients randomized to MED with or without CABG were analyzed in the surgical revascularization hypothesis. All Stratum C patients and Stratum B patients randomized to CABG with or without SVR were analyzed in the SVR hypothesis.

The clinical judgment of physicians and surgeons responsible for care of STICH-eligible patients determined the enrollment stratum offered for patient consent under the oversight of the ethics committee at each site. The primary ethical concern guiding equipoise for randomization was to offer patients treatment combinations judged to have similar long-term mortalities. Relative benefits of treatment combinations offered within each enrollment stratum might differ among STICH sites. The randomization process stratified treatment assignment by site. Whereas it is not possible to know reasons underlying enrolling physician judgment used to choose the enrollment stratum, it is possible to map the individual clinical characteristics of enrolled patients to provide evidence of the spectrum of disease severity produced by choices made.

### Creation of the Risk at Randomization Index

Three Duke Databank for Cardiovascular Diseases (DDCD) CAD prognosis publications (6-8) identified candidate variables used to create a Cox multivariable regression equation predictive of time to death in 821 STICH-eligible MED DDCD patients during 5.5 years of follow-up. Prognostic baseline clinical variables in descending order of importance for predicting time to death in this cohort over a mean follow-up of five years included age, renal disease (creatinine  $\geq 1.5$  mg/dL), heart failure, ejection fraction, Duke coronary artery disease index, mitral insufficiency, and cerebrovascular disease (Appendix 1). Individual STICH patient baseline values for these variables entered into the multivariable equation

produced probability of death for each STICH patient that ranged from 0.18387 to 0.99949. STICH patients were placed in ascending order of their individual prediction of death. All predicted death estimates summed to 1001.648 corresponded to a five-year mortality of 0.469, assuming only MED not surgical treatment for the 2,136 STICH randomized patients. Using least squares analysis to minimize differences between groups, 32 groups were created to each share 1/32 of total predicted deaths as close as was possible with the constraint that the exact risk of each patient had to be assigned to only one group. The numbers of patients in groups 1-32 decreased as the predicted risk increased except for groups 31 and 32 that each had 32 patients (Figure 1). Total predicted deaths per group ranged from a minimum of 30.89 (RAR 31) to a maximum of 31.67 (RAR 28). The average group mortality ranged from 0.228 in group 1 to 0.989 in group 32. By assigning the average group mortality to each patient in each group, every patient carries a marker of their predicted risk in the context of the entire randomized population. The RAR of each patient provides a tool to evaluate the impact of clinical site location and enrollment performance on the spectrum of risk of all patients enrolled. Creation of the RAR index only ordered patients by baseline risk without disrupting randomized status within RAR groups. Therefore, each patient may be considered to be interchangeable couriers of equivalent baseline risk when compared by clinical site characteristics within their RAR grouping.

### **Analysis of Clinical Site Location and Performance on Risk Spectrum of Enrolled Patients**

The RAR profiles of two dichotomous separations of the 2,136 STICH total patients were compared for the numbers of patients that shared each RAR interval. The RAR profile of 1,057 patients enrolled by the 13 highest-enrolling sites was compared to the RAR profile of 1,078 patients from the 114 lowest-enrolling sites. To evaluate the influence of enrolling country on risk distribution of patients, a mean patient weighted RAR was calculated to reflect the average risk of all patients enrolled in each country. The sum resulting from multiplying the RAR index number of each patient enrolled in each country when divided by total number of all enrolled patients in each country provided the mean patient weighted RAR for that country. Patients from three countries (Belgium, Greece, Turkey) enrolling only one patient were grouped with the four patients enrolled in Malaysia to provide a single patient-weighted mean RAR for the four lowest enrolling countries. Arranging countries in ascending order of patient-weighted mean RAR identified Sweden (22 patients) to separate 7 countries with lower mean RAR that enrolled 1,057 patients from the 18 countries with higher mean RAR that enrolled 1,057 patients. The 7 lower-risk patients from Sweden added to the 1,057 patients from lower-risk countries brought the number of patients in the lower-risk country grouping to 1,064. The 15 higher-risk patients added to the 1,057 patients from higher-risk countries brought the total to 1,072 patients in the higher-risk cohort. In RAR zones 1-12, clinical sites with a lower mean RAR enrolled most patients. In RAR zones 22-32, countries with higher mean RAR enrolled most patients. In RAR zones 13-21, both country groupings enrolled similar patients. For each of these three RAR zone groupings, baseline clinical characteristics of patients from countries tending to enroll high-risk patients were compared to patients from countries tending to enroll low-risk patients.

## **RESULTS**

### **Patient Enrollment**

Between July 24, 2002 and May 5, 2007, 127 sites in 26 countries randomized 2,136 patients. Based on clinical decisions and protocol guidelines, 1,277 (60%) of 2,136 randomized patients (Stratum A + Stratum B) were eligible for MED alone and 1,075 (50%) of the 2,136 patients (Stratum B + Stratum C) were SVR eligible (Figure 2). A relatively smooth profile of decline of number of patients was observed throughout the RAR risk

spectrum (Figure 3). The median age steadily increased from 48 years to 74 years, and the median ejection fraction decreased from 0.30 to 0.20 over the 32 RAR groups.

The same RAR group was shared by 890 (84%) of the 1,057 patients from higher enrolling sites and 82% of 1,078 patients from lower enrolling sites. The 167 unmatched patients from the higher enrolling sites were predominantly lower risk patients. The 189 unmatched patients from the lower enrolling sites were predominantly higher-risk patients. There was no consistent relationship between total patient enrollment in each country and the patient-weighted mean RAR index by enrolling country (Figure 4). Poland and Russia dominated enrollment of lower-risk patients, and the United States and Canada led countries enrolling higher-risk patients. RAR profiles showed 807 (76%) of 1,064 patients from countries enrolling lower-risk patients matched 807 (75%) of 1,072 patients from countries enrolling higher-risk patients (Figure 5). However, the RAR profile identified three patterns of unmatched patients. In RAR groups 1-12, all unmatched patients were from countries enrolling lower-risk patients. RAR groups 13-21 showed similar numbers of unmatched patients from both country groupings. In RAR groups 22-32, all unmatched patients were from countries enrolling higher-risk patients. Baseline clinical characteristics tabulated for these three RAR patient groupings demonstrated statistically significant differences between patients in the lowest and highest RAR groupings with less difference among patients in RAR groups 13-21 (Table 1). However, the magnitude of statistically significant differences of patients related to country of enrollment appeared sufficiently small to be of little clinical significance in application of STICH results to care decisions for future ischemic cardiomyopathy patients.

## DISCUSSION

The Northern New England Cardiovascular Disease Study Group identified patients with poor LV function to have the greatest inconsistency of CABG use compared to guideline recommendations (9). STICH investigators experienced this divergence of opinion when patients who were STICH-eligible by protocol were presented at investigator meetings. Difference in zones of equipoise among physicians responsible for enrolling patients at individual sites combined to produce a smooth risk profile for the entire STICH population. The four-fold decline in numbers of patients over the 32 RAR groups created with equivalent predicted deaths produced a STICH patient risk profile familiar to clinicians. Ischemic cardiomyopathy as the highest-risk subset of all chronic CAD diagnoses includes a rapidly declining number of patients in each RAR group that reflects the highest-risk tail of a normal distribution of all CAD patients truncated by an EF  $\leq 0.35$  applied as a STICH entry criterion.

Enrollment performance varied greatly among sites and was strongly related to baseline clinical risk. The highest enrolling sites tended to enroll more low-risk patients. The lowest-enrolling sites tended to enroll more high-risk patients. The patient risk zone that corresponded to the equipoise zone of the responsible physician appeared to be the strongest determinant of baseline risk of enrolled patients at a site. The profile of the RAR patient distribution suggests small differences in equipoise windows positioned toward the low- or high-risk direction could produce substantial changes in numbers of patients considered eligible for randomization. Many more low-risk patients than high-risk patients were available if sites were willing to randomize them. However, RAR data confirm that 85% of enrolled patients had similar predicted risk even when clinical sites were grouped by those sites with lowest or highest enrollment performance. Moreover, both site performance groupings enrolled some patients into each RAR group.



Country differences in population demographics also may influence the numbers of lower high-risk patients presenting with STICH eligibility criteria. The World Health Organization reports an age standardized mortality rate for cardiovascular disease per 100,000 population to be 324 for Poland and 688 for Russia, the two highest-enrolling countries of the 7 countries tending to enroll low-risk patients in contrast to rates of 188 for the United States and 141 for Canada, the two highest-enrolling of the 18 countries tending to enroll more high-risk patients (10). Country demographics, health care availability, and equipoise differences among physicians most likely contributed in varying degrees to produce observed variation in patient risk among enrolling countries. However, sufficient similarity of patient baseline risk among all enrolling countries has been demonstrated to suggest STICH enrollment of patients in 26 countries should enhance generalization of STICH trial results to a broad spectrum of future ischemic cardiomyopathy patients throughout the world.

Use of a multivariable equation to stratify baseline risks of randomized patients into low-, middle-, or high-risk groups was first reported by the Veterans Administration Cooperative Study Group in 1981 (11). Treatment effect was confined to only the high-risk cohort. This early work changed patient selection criteria for CABG from low- to high-risk patients and set the precedent for modeling of baseline characteristics of randomized patients to define their relative risk. The current work represents the first individualized patient-based RAR index used to describe the full risk spectrum of a randomized trial cohort.

## Limitations

Predicted mortality of STICH-eligible patients in the DDCD may not accurately reflect the mortality of STICH patients treated in other health care settings. However, the RAR grouping order of patients was independent of the numeric accuracy of the risk prediction.

RAR development included site-reported ejection fraction as the only assessment of LV function. Core laboratory assessments of global and regional LV function reported without knowledge of patient treatment will be used to assess the generalizability of STICH results based on the severity of abnormality represented by the SVR hypothesis patients. Many patient-specific clinical variables that influence decisions for cardiac surgery, such as conduit availability and suitability of coronary arteries for CABG, were not reported on STICH patients. However, all patients randomized were assessed as eligible for CABG or CABG with SVR by STICH surgeons at each clinical site.

## Conclusions

The RAR index methodology provides a clinically useful quantitation of the baseline risk spectrum of patients. Clinical site location introduced greater variation on the profile of baseline clinical risk than enrollment performance. However, 76% of patients shared the same risk profile when grouped by randomization in countries that enrolled more high- or low-risk patients. Differences in the baseline characteristics causing site-related differences in the RAR profile were modest and unlikely to hinder generalizability of STICH results. Diversity among patients randomized by 127 STICH clinical sites appears more likely to enhance generalizability of STICH trial results.

## Acknowledgments

The authors acknowledge with gratitude the support of Vanessa Moore in preparing the text of this article and of Kerry Bassett in the design and production of the graphic art essential to illustrate the message of this report.

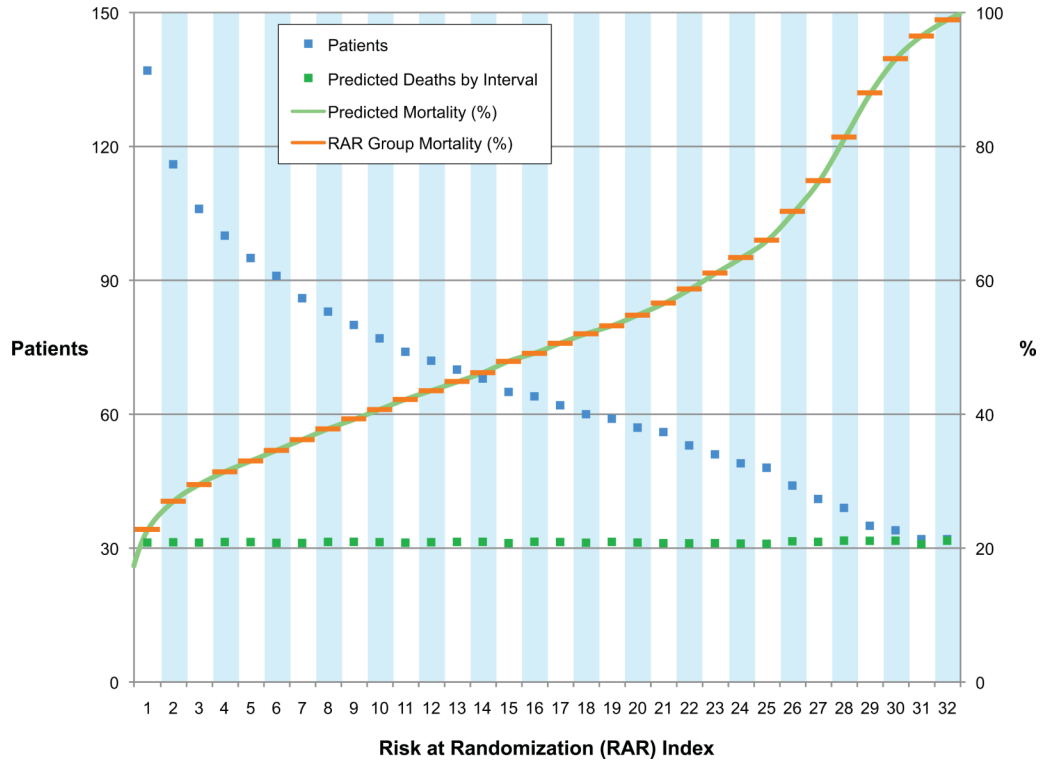
Supported by grant U01-HL69015 from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland

## ABBREVIATIONS

<b>CABG</b>	Coronary artery bypass grafting
<b>CAD</b>	Coronary artery disease
<b>DDCD</b>	Duke Databank for Cardiovascular Diseases
<b>LV</b>	Left ventricular
<b>MED</b>	Evidence-based medical therapy
<b>RAR</b>	Risk at randomization
<b>STICH</b>	Surgical Treatment for Ischemic Heart Failure trial
<b>SVR</b>	Surgical ventricular reconstruction

## References

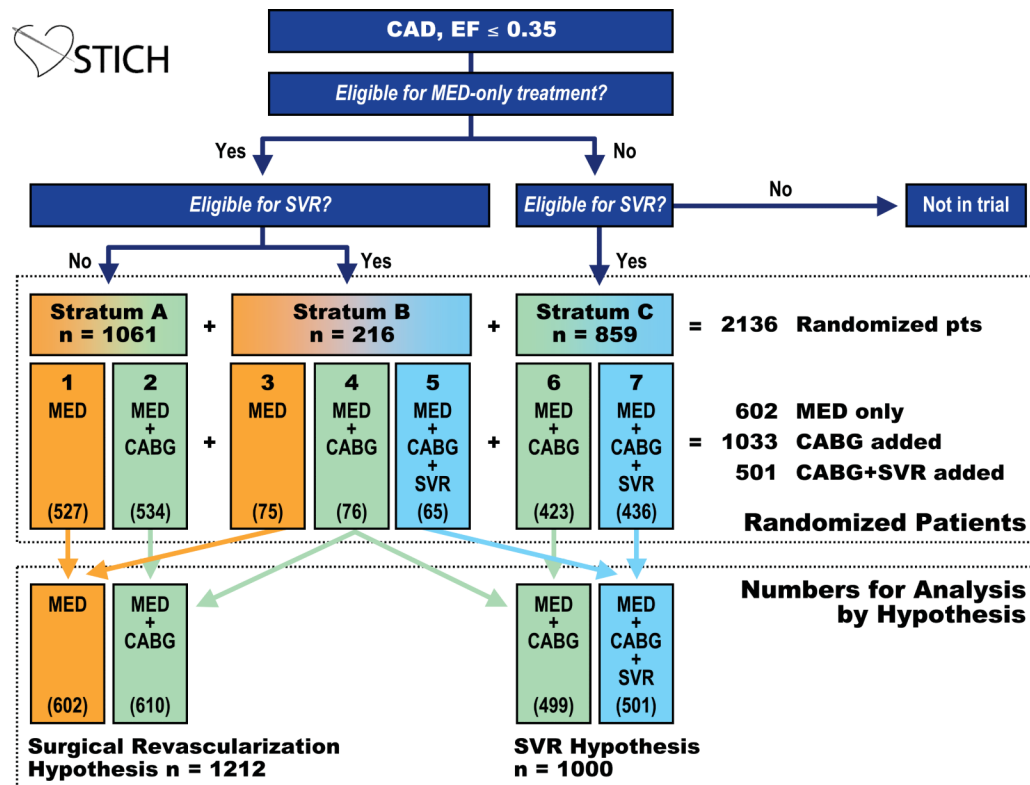
1. Velazquez EJ, Lee KL, O'Connor CM, Oh JK, Bonow RO, Pohost GM, Feldman AM, Mark DB, Panza JA, Sopko G, Rouleau JL, Jones RH. The rationale and design of the Surgical Treatment for Ischemic Heart Failure (STICH) Trial. *J Thorac Cardiovasc Surg.* 2007; 134:1540–7. [PubMed: 18023680]
2. Jones RH, Velazquez EJ, Michler RE, et al. Coronary bypass surgery with or without surgical ventricular reconstruction. *N Engl J Med.* 2009; 360:1705–17. [PubMed: 19329820]
3. Jones RH. Correspondence – Surgical Ventricular Reconstruction. *NEJM.* 2009; 361:529–532.
4. Blair JEA, Zannad F, Konstam MA, et al. Continental differences in clinical characteristics, management, and outcomes in patients with worsening heart failure. *J Am Coll Cardiol.* 2008; 52:1640–8. [PubMed: 18992654]
5. Glickman SW, McHutchison JG, Peterson ED, et al. Ethical and scientific implications of the globalization of clinical research. *N Engl J Med.* 2009; 360:816–23. [PubMed: 19228627]
6. Califf RM, Harrell FE Jr, Lee KL, et al. The evolution of medical and surgical therapy for coronary artery disease: a 15-year perspective. *JAMA.* 1989; 261:2077–86. [PubMed: 2784512]
7. Jones RH, Kesler KK, Phillips HR III, et al. Long-term survival benefit of coronary artery bypass grafting and percutaneous transluminal angioplasty in patients with coronary artery disease. *J Thorac Cardiovasc Surg.* 1996; 111:1013–1025. [PubMed: 8622299]
8. Smith PK, Califf RM, Tuttle RH, et al. Selection of surgical or percutaneous coronary intervention provides differential longevity benefit. *Ann Thorac Surg.* 2006; 82:1402–29.
9. O'Connor GT, Olmstead EM, Nugent WC, et al. for the Northern New England Cardiovascular Disease Study Group. Appropriateness of coronary artery bypass graft surgery performed in Northern New England. *J Am Coll Cardiol.* 2008; 51:2323–2328. [PubMed: 18549917]
10. World Health Organization. <http://www.who.int/whosis/indicators/compendium/2009/net>
11. Detre K, Peduzzi P, Murphy M, et al. for the Veterans Administration Cooperative Study Group for Surgery for Coronary Arterial Occlusive Disease. Effect of bypass surgery on survival in patients in low- and high-risk subgroups delineated by the use of simple clinical variables. *Circulation.* 1981; 63:1329–1338. [PubMed: 6971716]



**Figure 1. Risk at Randomization Development**

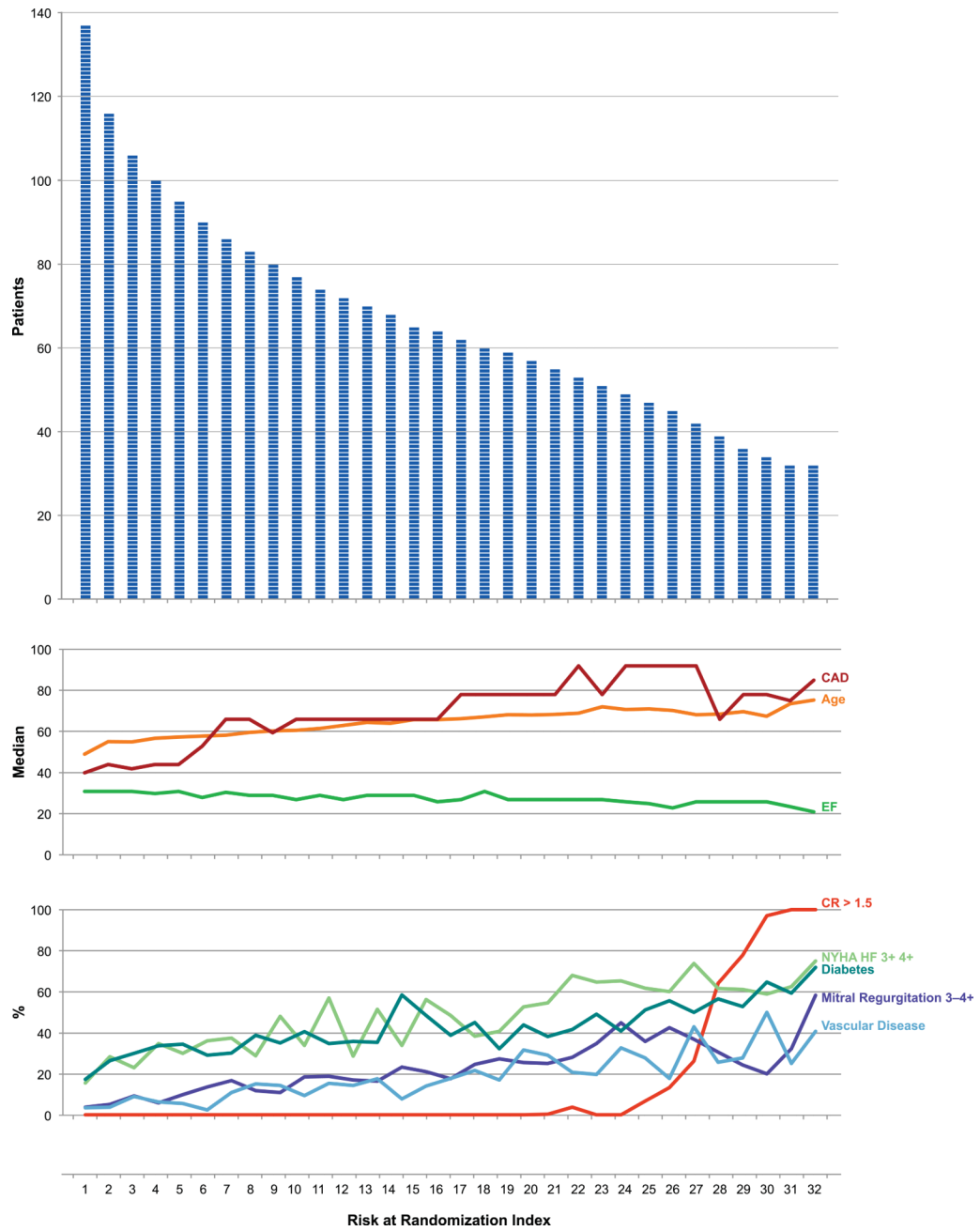
The continuous green line represents the 2136 data points of individual patient mortality predicted from the multivariable model. The interrupted green lines represent total predicted deaths for all patients in each RAR group. The interrupted blue lines depict the total number of patients in each RAR group. The interrupted orange line represents the RAR group mortality assigned to each patient in each RAR group.





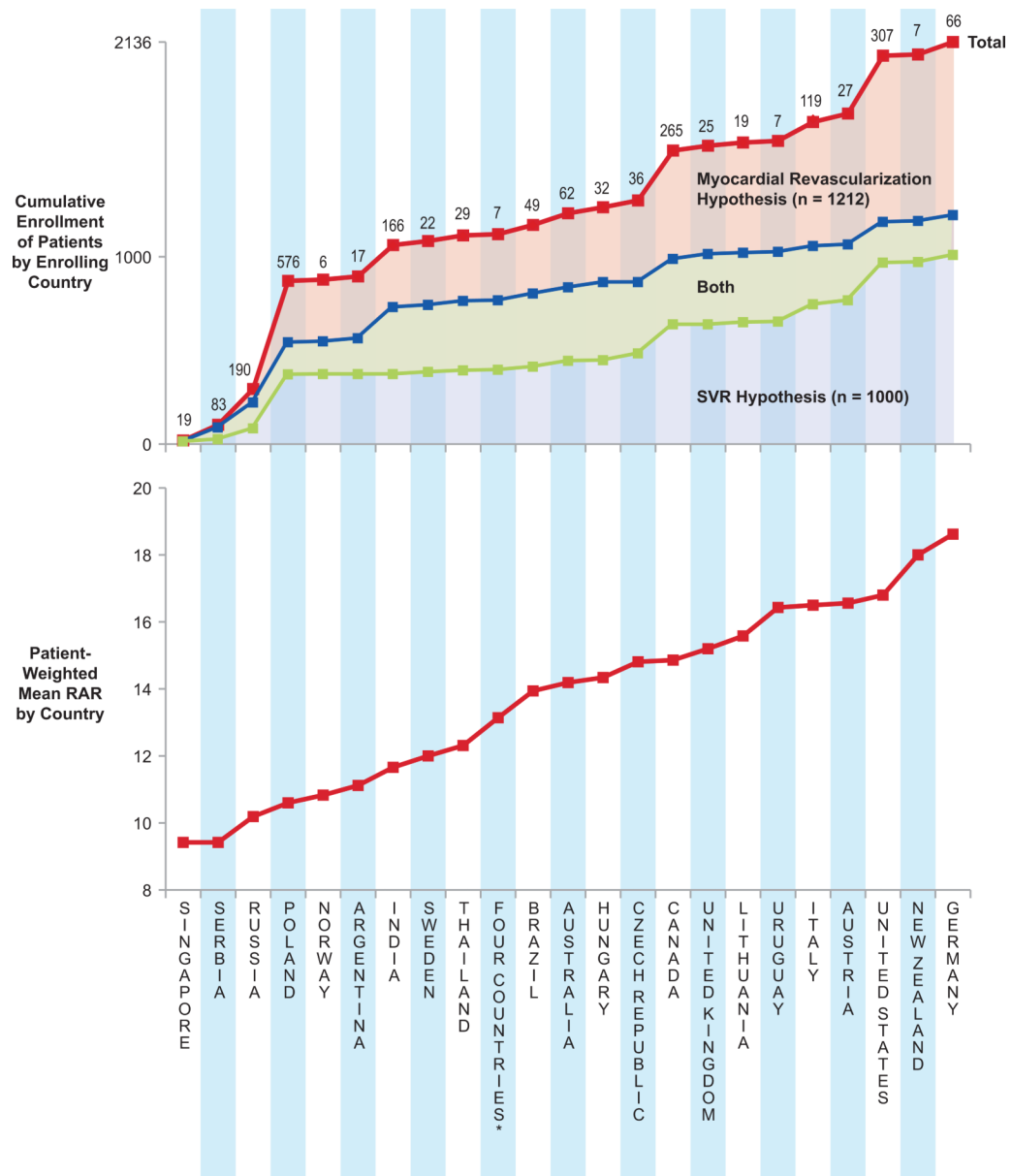
**Figure 2. Schema of Patient Enrollment**

This schematic depiction of the STICH trial design includes numbers of patients randomized to each stratum. Treatment assignments were MED only for 602 patients, CABG added to MED for 1033 patients, and CABG and SVR added to MED for 501 patients. The 76 Stratum B patients assigned to MED + CABG treatment are used to address both the surgical revascularization hypothesis and the SVR hypothesis. Therefore, 2136 randomized patients provided 2212 patients with potential for primary endpoints for analysis of either or both hypotheses.



**Figure 3. Patients and Corresponding Clinical Characteristics by RAR**

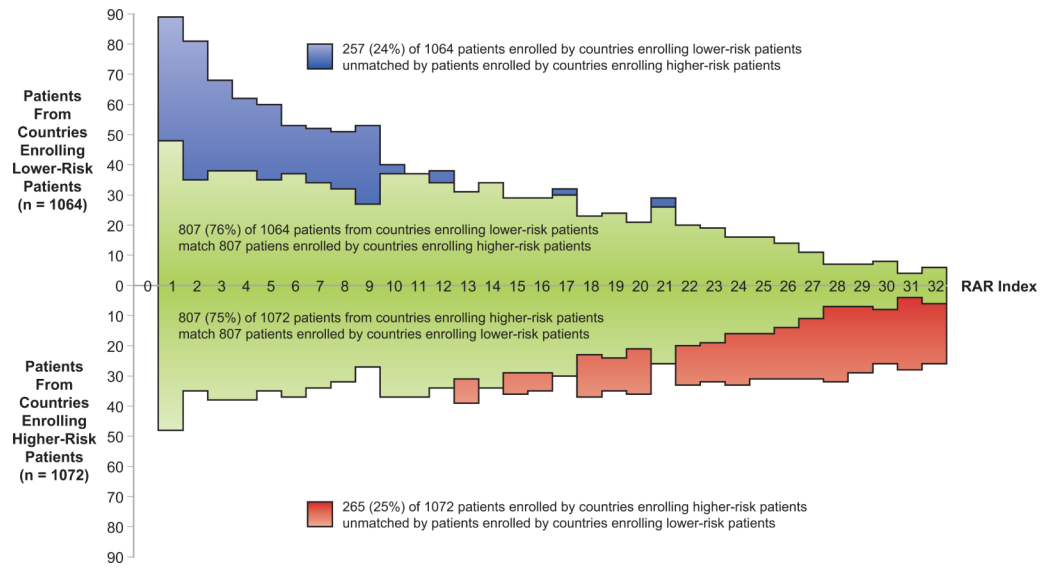
This histogram depicts the individual patients at each RAR interval. The corresponding median age, ejection fraction, and Duke CAD index for each RAR group and the percent of patients in each RAR group with baseline creatinine  $\geq 1.5$ ,  $\geq$  NYHF class III, diabetes, mitral regurgitation  $\geq 3+$ , and vascular disease provided a comprehensive summary of the components of risk for the entire population.



\*Four countries with few patients (Belgium 1, Greece 1, Turkey 1, Malaysia 4) combined

**Figure 4. Country Enrollment by Enrolled Cohort Mean RAR**

Countries ranked by the patient-weighted RAR mean show a two-fold range of difference from the lowest 9.4 (Singapore) to highest 18.6 (Germany) patient-weighted RAR mean. The numbers of patients enrolled in each country and cumulative patient enrollment by increasing RAR show the diversity of enrollment performance by country throughout the spectrum of patient risk of enrolling countries. The patient-weighted mean RAR index was calculated by multiplying the numbers of patients enrolled by each country by the index number of each RAR interval. The sum of these products is divided by the total numbers of patients enrolled by that country to produce the patient-weighted mean RAR.



**Figure 5. RAR Profile by Country Mean Patient Risk**

A total of 1,614 (76%) of 2,136 patients shared the same RAR grouping despite intentional creation of two cohorts to maximize imbalance among countries based on the average baseline risk of patients randomized. RAR groups 1-12 included more patients from countries with low patient-weighted mean RAR, and RAR groups 22-32 included more patients from countries with high patient-weighted mean RAR. RAR groups 10-21 reflected the most even distribution of patients from the two country groupings of sites. The baseline characteristics of patients in these two groupings are compared in Table 1.

