SAN ANTONIO BREAST CANCER SYMPOSIUM®
UT Health San Antonio Mays Cancer Center
AACR American Association for Cancer Research
DECEMBER 9–12, 2025
HENRY B. GONZALEZ CONVENTION CENTER • SAN ANTONIO, TX

Multimodal Artificial Intelligence (AI) Models Integrating Image, Clinical, and Molecular Data for Predicting Early and Late Breast Cancer Recurrence in TAILORx

Joseph A. Sparano[1], Norsang Lama[2], Robert J. Gray[3], Md Ashequr Rahman[2], Victoria Wang[3], Della F. Makower[4], Yating Cheng[2], Kathy S. Albain[5], Ming Chen[2], Daniel F. Hayes[6], Anthony Helmstetter[2], Charles E. Geyer, Jr.[7], Casey Bales[2], Elizabeth C. Dees[8], Matthew P. Goetz[9], John A. Olson, Jr.[10], Sunil S. Badve[11], Thomas J. Saphner[12], Timothy J. Whelan[13], Virginia G. Kaklamani[14], Matthew Oberley[2], Milan Radovich[2], David Spetzler[2], Eleftherios P. Mamounas[15,16], Norman Wolmark[16], George W. Sledge[2].

Icahn School of Medicine at Mount Sinai, Tisch Cancer Institute, NY, NY[1], Caris Life Sciences, Irving, TX[2], ECOG-ACRIN Biostatistical Center, Dana-Farber Cancer Institute, Boston, MA[3], Montefiore Einstein Comprehensive Cancer Center, Bronx, NY[4], Loyola University Medical Center, Maywood, IL[5], University of Michigan, Ann Arbor, MI[6], University of Pittsburgh, Pittsburgh, PA[7], University of North Carolina, Chapel Hill, NC[8], Mayo Clinic, Rochester, MN[9], Washington University School of Medicine, St. Louis, MO[10], Winship Cancer Institute, Emory University, Atlanta, GA[11], Aurora Medical Center, Two Rivers, WI[12], McMaster University, Hamilton, CAN[13], University of Texas Health, San Antonio, TX14, AdventHealth Cancer Institute, Orlando, FL[15], National Surgical Adjuvant Breast and Bowel Project, Pittsburgh, PA[16]

Mount Sinai Tisch Cancer Center

Canadian Cancer Trials Group · Alliance for Clinical Trials in Oncology · Federally Funded Research · ECOG-ACRIN cancer research group · Reshaping the future of patient care · National Cancer Institute · SWOG Leading cancer research. Together. · NRG Oncology

CARIS LIFE SCIENCES

Disclosure Information

Joseph A. Sparano, MD

I have the following relevant financial relationships to disclose:

Employee of: Icahn School of Medicine at Mount Sinai, Tisch Cancer Institute, New York, NY

Consultant for: Astra Zeneca, Genentech/Roche, Novartis, Pfizer,

Speaker's Bureau for: None

Grant/Research support from: National Cancer Institute, Department of Defense, Breast Cancer Research Foundation

Stockholder in: None

Honoraria from: None

Background: Trial Assigning Individualized Options for Treatment (TAILORx)

- **Rationale:**
  - Oncotype DX (ODX) 21-gene RS was known to be prognostic for distant recurrence & predictive of chemotherapy benefit in ER+, HER2-neg, node-negative breast cancer
- **Design:**
  - TAILORx was designed to address uncertainty of CT benefit in mid-range RS 11-25, and prospectively confirm the favorable prognosis with ET alone for a very low RS 0-10
- **Results:**
  - Adjuvant ET is non-inferior to CET for 69% with a 21-gene RS 11–25 (primary endpoint)
    - ❑ Exploratory analysis: age ≤50 years had some CT benefit that varied with RS and clinical risk
  - RS 0–10 is associated with low DR rates with ET alone
  - Integration of RS with clinicopathologic features using the RSClin tool was more prognostic for distant recurrence than either used alone

Paik et al. NEJM 2004; Paik et al. JCO 2006; Sparano et al. NEJM 2015; Sparano et al. NEJM 2018; Sparano et al. NEJM 2019; Sparano et al. NEJM Evidence 2024

Objective and Methods

- **Objective:**
  - Develop a diagnostic test with better distant recurrence (DR) risk prognostication than the Oncotype DX (ODX) 21-gene recurrence score (RS)
- **Sample Analysis:**
  - Primary FFPE tumor samples analyzed for whole transcriptome sequencing (WTS) and whole exome sequencing (WES) by Caris MI Tumor Seek–Hybrid
  - H&E images were digitized at 40X and analyzed at 10x by deep learning
- **Statistical Analysis:**
  - Models reflecting clinical (C), expanded molecular (M+), and pathomic imaging (I) features
  - Model training & 5-fold nested cross validation (60/20/20) for distant recurrence
  - Model performance was compared with the actual ODX RS
  - Assessed using the concordance index (C-index), hazard ratios (HR), & log-rank tests
  - Best performing models independently evaluated in holdout validation set

Methods: REMARK Summary for Training/5-fold Cross-Validation and Holdout Validation Sets

SAN ANTONIO BREAST CANCER SYMPOSIUM®

UT Health San Antonio Mays Cancer Center | AACR American Association for Cancer Research

- **10,273 enrolled in TAILORx**
  - 8,062 (78%) with FFPE tissue blocks
  - 6,530 (64%) with evaluable H&E 40x WSI generated by Pramana scanner
- **4,462 (43%) cases with WSI & WTS**
  - **2,808 training/5-fold CV (63%)**
    - **2482 with WES used for exploratory modeling**
  - **1,621 holdout validation (37%)**

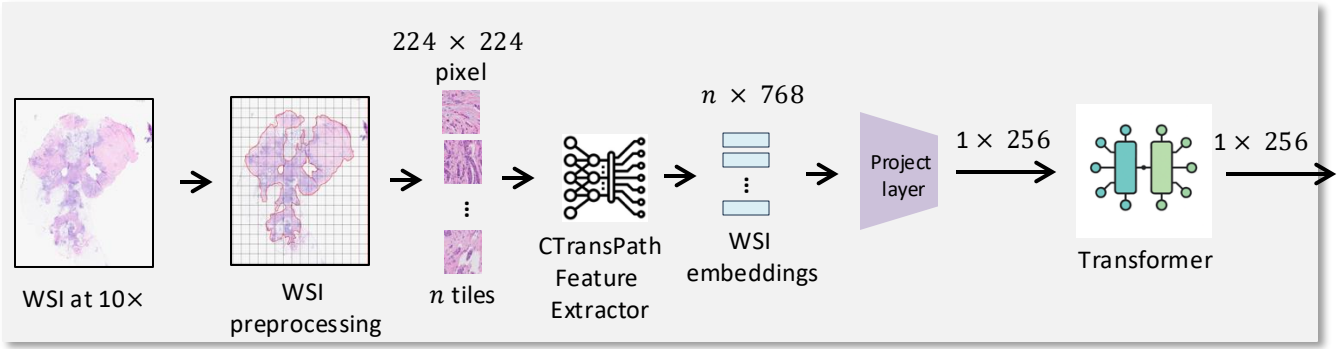
Training set: ECOG-ACRIN, Alliance (ACOSOG, CALGB, NCCTG) & CCTG
Holdout set : enrolled through SWOG, NRG (NSABP, RTOG, GOG) & CTSU

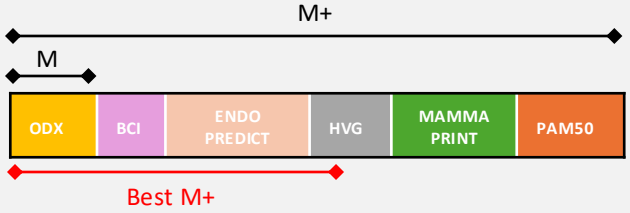
Clinicopathologic characteristics

- **Training & validation set *not* significantly different:**
  - Median age: 56 years (range 23-75 years)
  - Postmenopausal: 66.2%
  - Tumor size ≤ 2 cm: 72.4%
  - PR positive: 90.1%
  - Low clinical risk: 66.9%
- Statistically significant but clinically marginal differences:
  - High grade: 20.7% vs. 19.7% (p=0.004)
  - Low grade: 24.2% vs. 28.4%

- **Included vs. excluded *not* significantly different:**
  - Age, menopausal status, PR, surgery, arm
- Statistically significant but clinically marginal differences:
  - Tumor size ≤ 2 cm: 72.4% vs. 76.9% (p<0.001)
  - High grade: 19.7% vs. 17.2% (p=0.005)
  - Low grade: 25.8% vs. 26.7%
  - Low clinical risk: 66.9% vs. 71.9% (p<0.001)

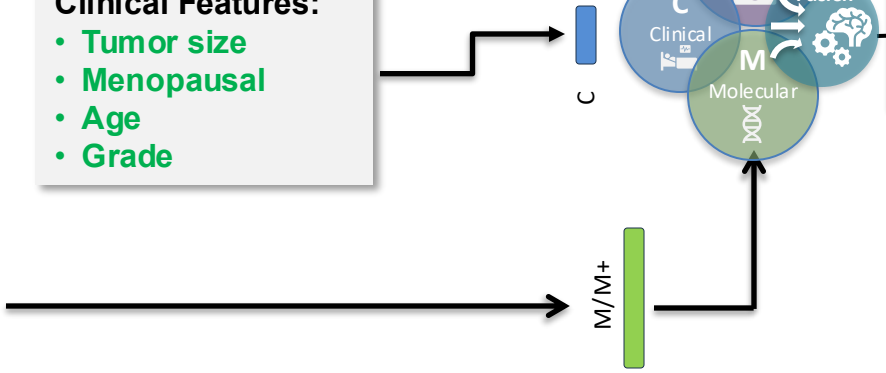
Multimodal Model Architecture





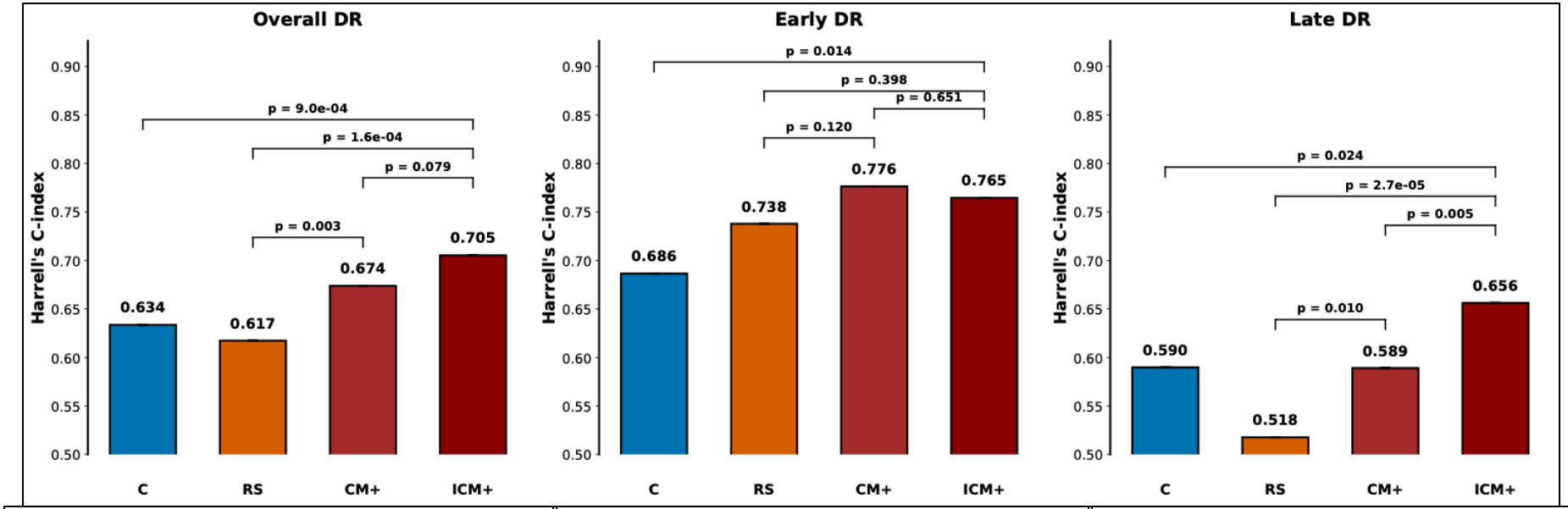
Different combinations of gene sets were tested for M+, best M+ consists of all ODX, BCI, EndoPredict, and some HVG genes

Clinical Features:
- Tumor size
- Menopausal
- Age
- Grade



WSI: Whole Slide Image; **HVG**: Highly Variable Gene; **MLP**: Multi-Layer Perceptron

Results: C-index for Optimal ICM+ and CM+ Models vs. Clinical Model and ODX RS in Training/5-fold CV Set



Overall DR at 15 Years

- ICM+ & CM+ model performance significantly better than C model & RS
- ICM+ model had highest C-index, but not significantly better than CM+

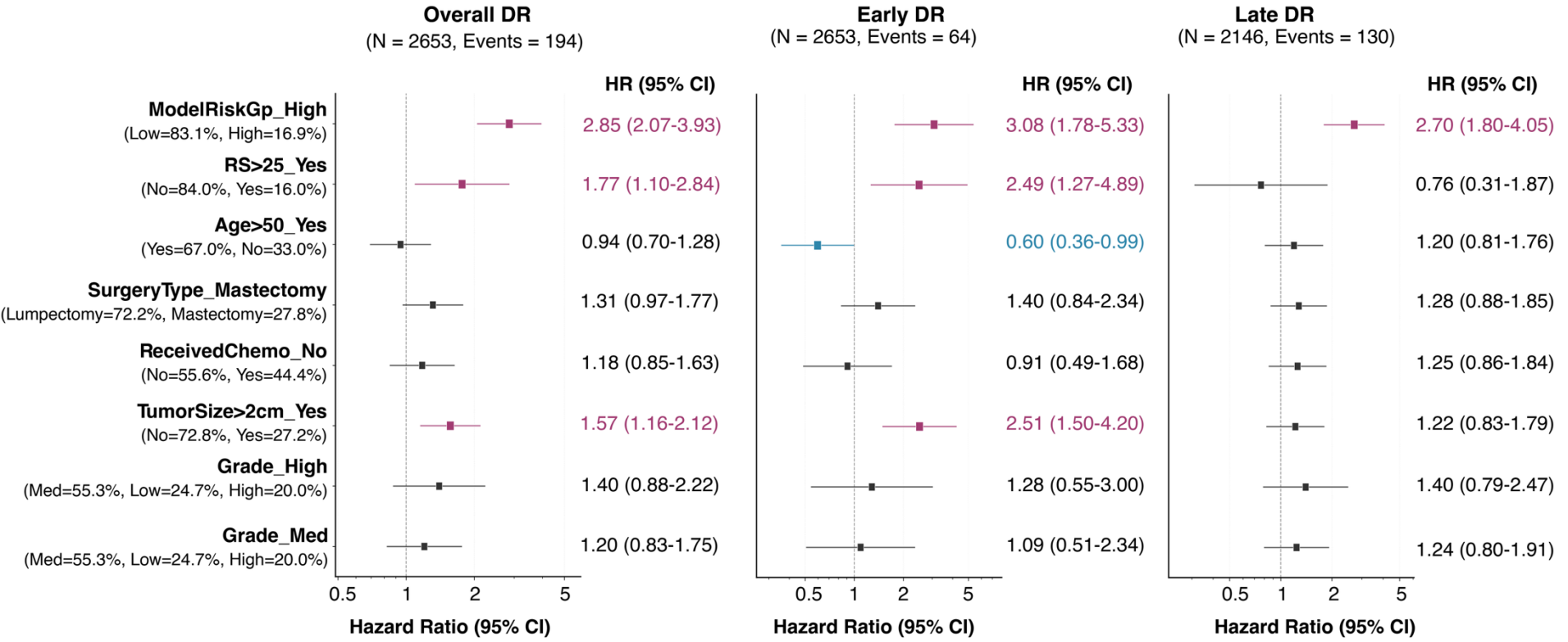
Early DR </=5 Years

- ICM+ and CM+ models exhibited highest C-index, better than C model
- ICM+ and CM+ models numerically but not significantly better than RS

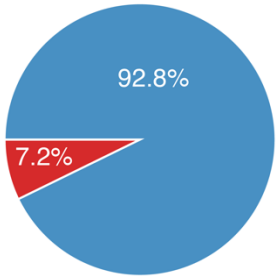
Late DR > 5 Years

- ICM+ model performance significantly better than C, RS, and CM+

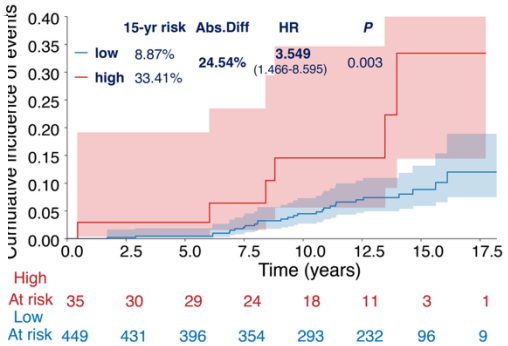
Results: Multivariate Analysis for Distant Recurrence Including the ICM+ Model and Other Covariates in the Training/5-fold CV Set



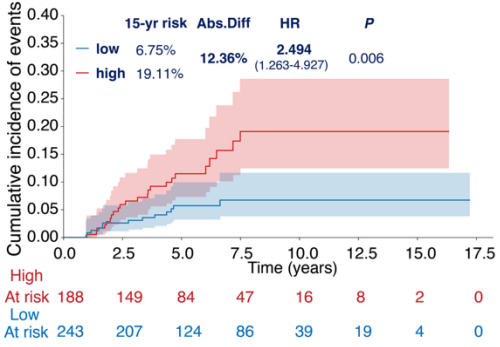
Results: Distribution of ICM+ Risk Groups by RS-based Treatment Arms and 15-year Distant Recurrence Risk in Training/5-fold CV Set

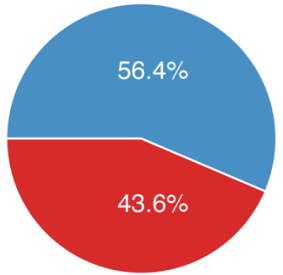


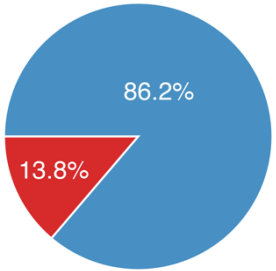
RS0-10: ET



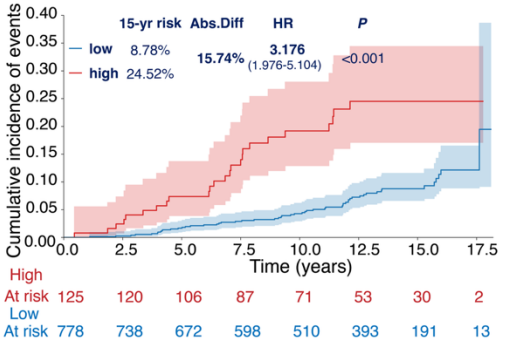
RS26-100: CET



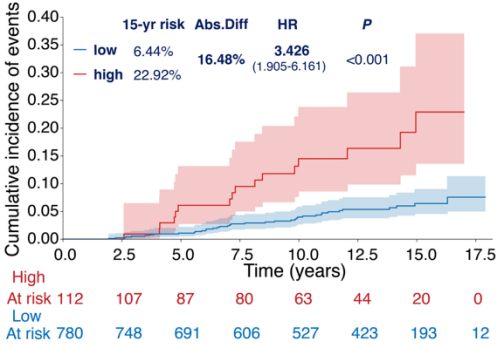


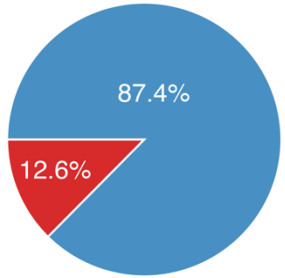


RS11-25: ET

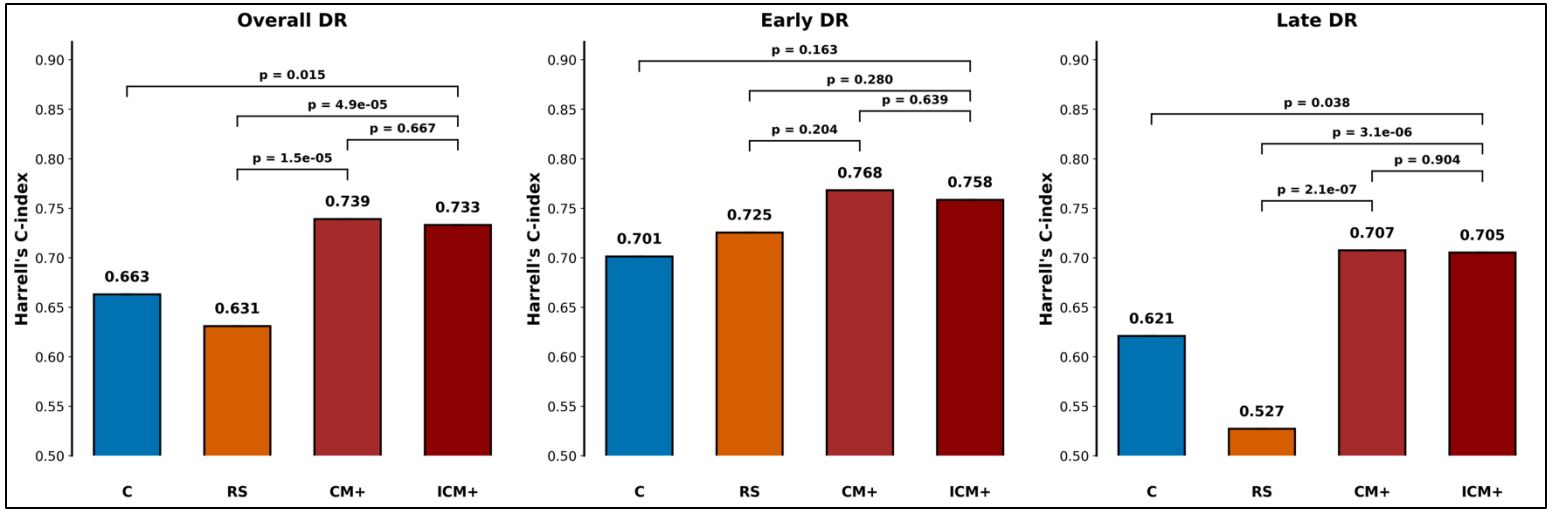


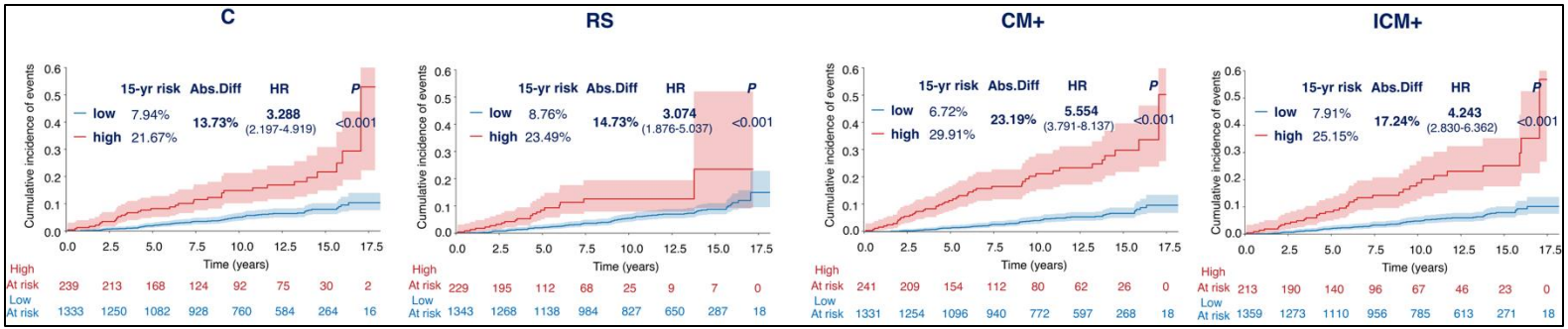
RS11-25: CET



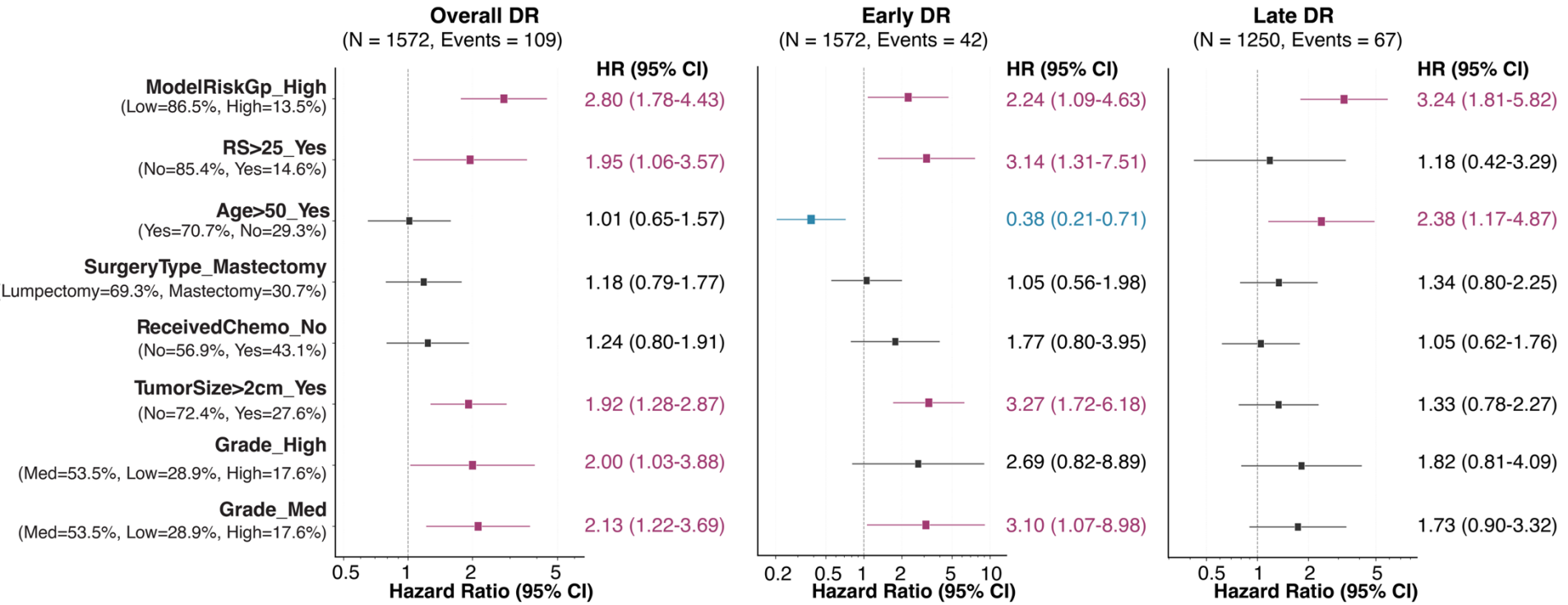


Results: C-index for Optimal ICM+ and CM+ Models vs. Clinical Model and ODX RS in Holdout Validation Set

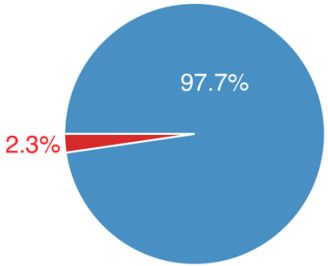




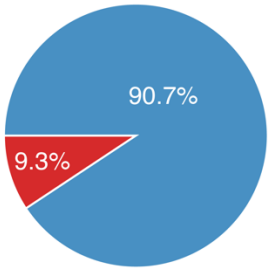
Results: Multivariate Analysis for Distant Recurrence Including the ICM+ Model and Other Covariates in the Holdout Validation Set



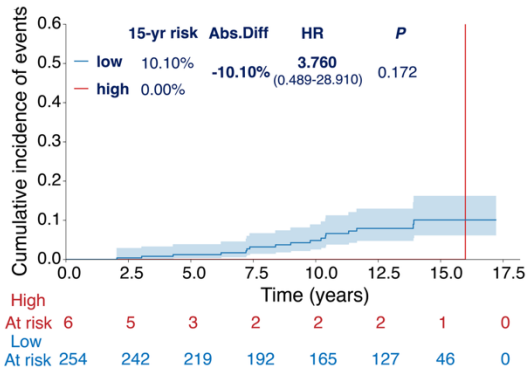
Results: Distribution of ICM+ Risk Groups by RS-based Treatment Arms and 15-year Distant Recurrence Risk in Holdout Validation Set



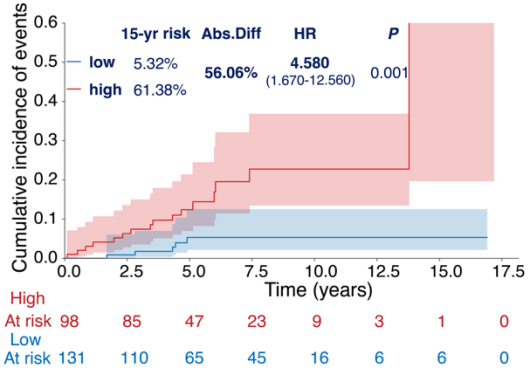
High Risk
Low Risk

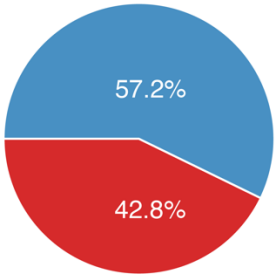


RS0-10: ET

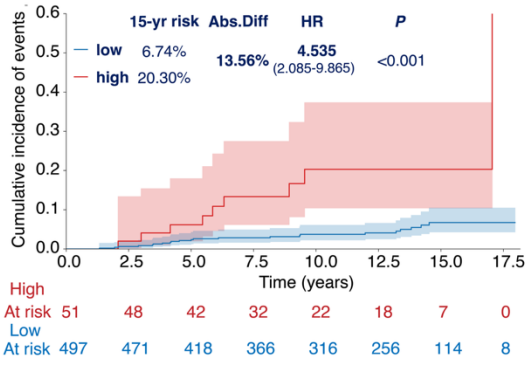


RS26-100: CET

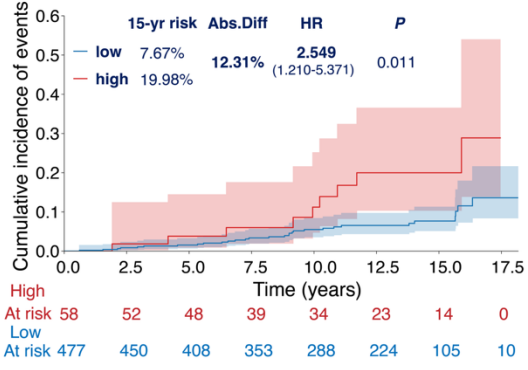




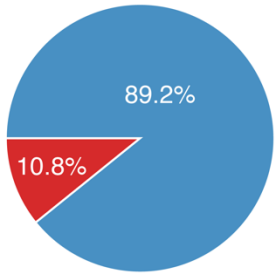
RS11-25: ET



RS11-25: CET



High Risk
Low Risk



Conclusions

- **Training/5-fold cross validation set:**
  - **C Model:** Clinical features provided some prognostic information
  - **M+ Model**: 42-gene molecular model including the 21-gene RS, BCI, and EndoPredict primarily drove prognostic stratification for early DR within 5 years
  - **I Model:** Pathomic features strengthened prognostic stratification for late DR > 5 years
  - **Multimodal ICM+ Model**:
    - ❑ Strongest prognostic performance for overall and late DR that were superior to the actual ODX RS
    - ❑ Provided statistically significant and clinically relevant prognostic stratification in the low (RS 0-25) and high (RS 26-100) genomic risk groups
    - ❑ Exhibited stronger performance than **CM+** for late recurrence
- **Holdout validation set:**
  - Superior prognostic performance of both the **ICM+** & **CM+** models were validated for overall and late DR
  - Both **ICM+** & **CM+** models outperformed the actual Oncotype DX RS used to guide therapy

Other TAILORx Presentations and EA1241

RF3-01: Clinical outcomes of invasive lobular carcinoma (ILC) versus non-lobular breast cancer (NLC) assessed by expert pathologists, an artificial intelligence (AI) CDH1 classifier, and AI-derived tumor microenvironment (TME) biomarkers in TAILORx.

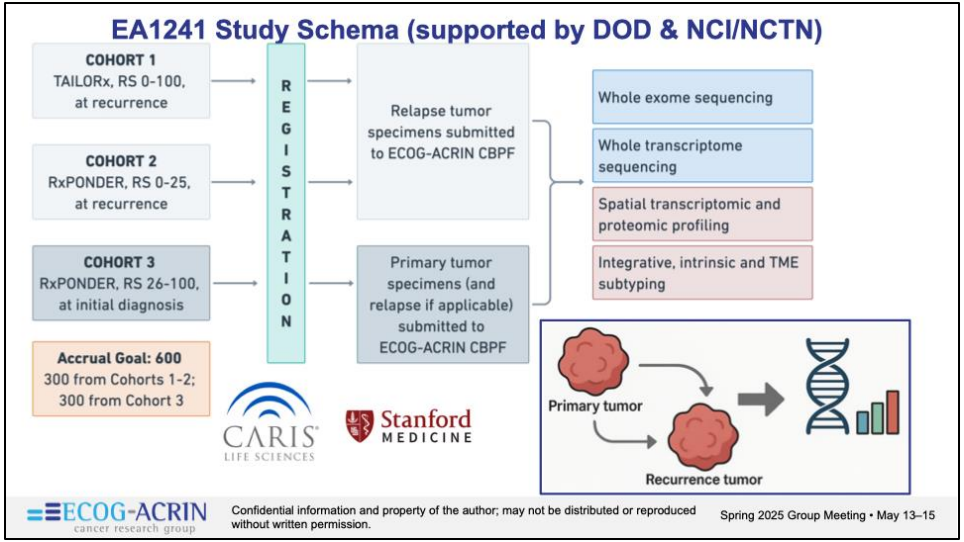
Roberto Salgado et al on behalf International Immuno-Oncology Biomarker Working Group and the TAILORx Investigators

RF3-07: A Multimodal-Multitask Deep Learning Model Trained in NSABP B-42 and Validated in TAILORx for Late Distant Recurrence Risk in HR+ Early Breast Cancer.

Eleftherios (Terry) Mamounas et al on behalf of NRG/NSABP and the TAILORx Investigators

Roberto Salgado, MD
IO Biomarker WG

Terry Mamounas, MD, MPH
NRG/NSABP



Kevin Kalinsky, MD
SWOG

Christina Curtis, PhD
Stanford Medicine

Rima Patel, MD
ECOG-ACRIN

Acknowledgements

- TAILORx trial volunteers
- Cancer advocacy community
- Health care & research professionals
- Funders
  - Breast Cancer Research Foundation
  - NIH/NCI: Award numbers: U10CA180820, U10CA180794, UG1CA189859, UG1CA233160, UG1CA233184, UG1CA233373, UG1CA232760, UG1CA233339, UG1CA233247, and UG1CA190140
  - U.S. Postal Service Stamp Fund