Refining Chemotherapy Decisions in Fertility-Preserving, Low-Risk ER+/HER2− Breast Cancer via ML Genomic Subgrouping

Wanru Guo[1], Curtis Tatsuoka[1]
Department of Biostatistics, University of Maryland, Baltimore, Baltimore, MD, USA[1]; University of Maryland Greenebaum Comprehensive Cancer Center, Baltimore, MD, USA[1]. wanru.guo@som.umaryland.edu

Background & Objective

- Invasive Ductal Carcinoma is the most common breast cancer subtype; ER+ patients typically receive endocrine therapy.
- Chemotherapy may offer added benefit for select high-risk cases.
- For women <50, treatment decisions should consider fertility and overtreatment.
- Low NPI patients are usually spared chemo, although some genetic subgroups may still benefit.
- **Goal:** Use a two-stage machine learning framework to:
  - Identify genomic signatures of young ER+/HER2− patients who benefit from combination chemo-endocrine therapy (chemo in addition to hormonal therapy).
  - Detect genomic subgroups with hidden chemo benefit (alongside hormonal therapy) in low-NPI patients.

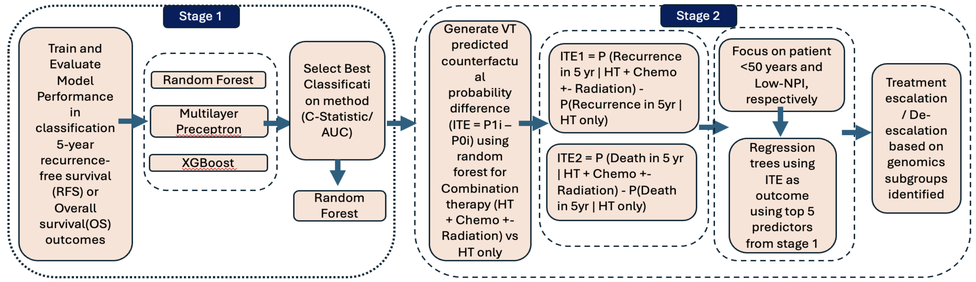
Methods

- **Dataset**: Longitudinal cohort of 423 IDC patients; all received adjuvant treatment post-surgery.
- **Predictors**: 19 clinical and genomics predictors - age, tumor stage, histology, laterality, cellularity, PR status, claudin-subtype, multi-omics clusters, TMB, mutation count, Nottingham Prognostic Index (NPI) (Note: Age was excluded when modeling the <50 group; NPI was excluded when modeling the low-NPI group)
- **Outcomes:** 5-year overall survival (OS) and recurrence-free survival (RFS), coded as binary endpoints (alive/dead; recurrence/no recurrence).
- **Virtual Twins Framework**:

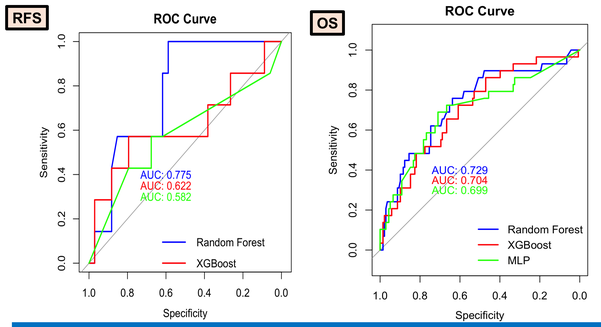
Stage 1: Compare classifiers (Random Forest, XGBoost, MLP) for OS/RFS prediction. Best-performing model used to estimate Individual Treatment Effects (ITE).

Stage 2: Train regression tree on ITEs using top 5 predictors to identify treatment-sensitive subgroups.

- **Focus**: Stratify by age (<50) and low NPI to identify groups benefiting from chemo-endocrine therapy.

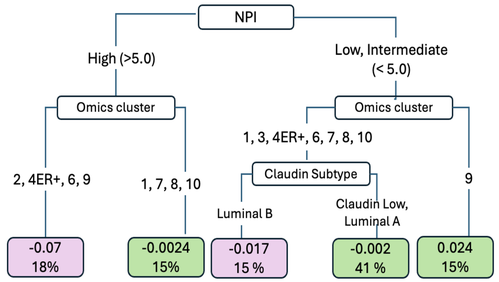


Stage 1 VT – Select Best Model in Prediction of RFS & OS



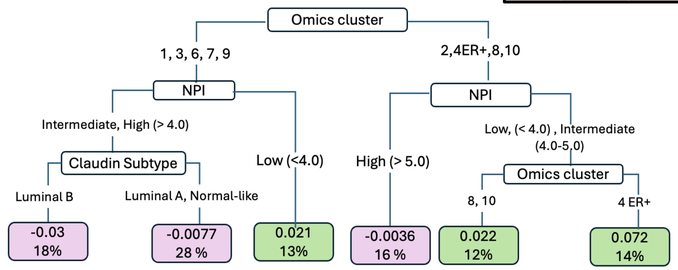
Stage 2 VT – 1. Genomic Treatment Sensitivity Subgroups in Fertility-Conscious Patients (Age < 50)

Age < 50 years OS



- **Highest OS benefit (ITE = –0.07, 18%):**
  - High NPI with 11q13/14 (*CCND1, PAK1, RSF1, EMSY*), 8p12, 8q, 20q amps
  - Plus TRG/TRA deletions, TCR disruption, CD8+ infiltration
  - Indicates proliferative, immune-inflamed, chemo-sensitive subtype
- Luminal B–like, Low NPI, **Modest Benefit (ITE = –0.017, 15%):**
  - Dual signature: moderate proliferation + immune activity
- **Intermediate benefit (ITE ≈ –0.0024, 15%):**
  - Genomic instability: 17q23, 20q, 5q loss, 8q gain, 10p/12p (*TTK, AURKB, CDC20, FOXM1, RAD51AP1*)
  - Plus 16p/16q, 1q/16q switches → subclonal heterogeneity
- **Minimal/no benefit (ITE = 0.024, 15–41%):**
  - Low NPI, 8q/20q alone, 1q gain, 16q loss, Luminal A / Claudin-low
  - Hormone-driven, genomically quiet, immune-cold tumors

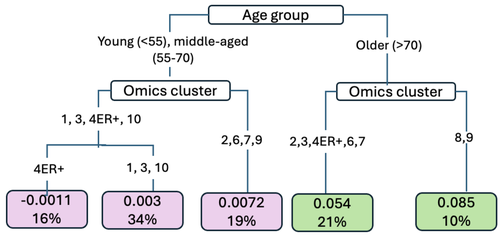
Age < 50 years RFS



- **Greatest RFS benefit (ITE = –0.03, 18%)**
  - Intermediate–high NPI (>4.0), Luminal B subtype
  - Genomic features: 17q23, 20q amplifications, 8p12, 8q, 16p gain/16q loss, 8q/20q gains
- **Moderate benefit (ITE = –0.0077, 28%)**
  - Same genomic context with Luminal A or Normal-like subtype
- **Additional benefit group (ITE = –0.0036, 16%)**
  - High NPI with amplifications at 11q13/14 (CCND1, PAK1, RSF1, EMSY); Genomic instability: 5q loss, 8q gain, 10p/12p gains (TTK, AURKB, CDC20, FOXM1, RAD51AP1), 1q gain/16q loss; High immune infiltration
- **Minimal benefit (ITE = 0.021–0.072, 12–14%)**
  - Low/intermediate NPI
  - Genomically quiet or modestly altered tumors (e.g., 8, 10; cluster 4 ER+)
  - TRG/TRA deletions with CD8+ infiltration showed the least benefit **(ITE = 0.072)**

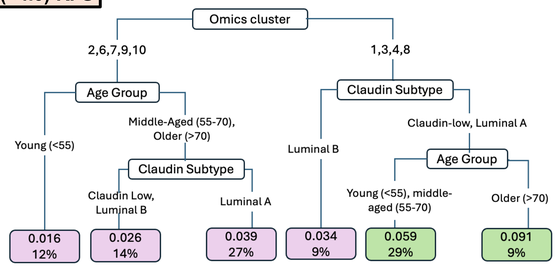
Stage 2 VT – 2. Genomic Treatment Sensitivity Subgroups in Low-NPI Patients (<4.0)

Low-NPI (<4.0) OS



- **Most OS benefit**
  - Young (<55) and middle-aged (55–70) with high CD8+ infiltration and complex genomic instability: **(ITE = –0.0011, 16%)**
  - 17q23 & 20q amplifications (1), 5q loss, 8q gain, 10p/12p gains (TTK, AURKB, CDC20, FOXM1, RAD51AP1) (10) **(ITE = 0.003, 34%)**
- **Modest benefit (ITE = 0.0072, 19%)**
  - Patients <70 with amplifications in - 11q13/14 (CCND1, PAK1, RSF1, EMSY), 8q12, 16p gain/16q loss, 8q, and 8q/20q
- **Least benefit (ITE = 0.085, 10%)**
  - Older patients (>70 years) with 1q gain and 16q loss or 8q/20q gains;
  - Reflects hormone-driven, genomically quiet Luminal A-like biology

Low-NPI (<4.0) RFS



- **Most RFS benefit (12–14%)**:
  - Young patients (<55) with amplifications in 11q13/14 (CCND1, PAK1, RSF1, EMSY), 8p12, 8q, 20q, 16p gain/16q loss, and mitotic/DNA repair genes (TTK, AURKB, CDC20, FOXM1, RAD51AP1) **(ITE = 0.016, 12%);**
  - Claudin-low and Luminal B subtypes in middle-aged/older patients also show moderate benefit **(ITE = 0.026, 14%)**
- **Moderate benefit (27%)**:
  - Luminal A tumors among middle-aged/older patients within the same genomic subgroup **(ITE = 0.039, 27%)**
- **Least benefit (9–29%)**:
  - Older patients (>70) with Luminal A or Claudin-low subtypes, genetically quiet tumors or 1q gain/16q loss **(ITE = 0.091, 9%).**
  - Younger (<70) in this subgroup show slightly better outcome for combination therapy **(ITE = 0.059, 29%).**

Clinical Implications of Findings

Age < 50 years

- **Escalation recommended**:
  - **High NPI**, Clusters **2, 4ER+, 6, 9** → strong OS benefit
  - **Intermediate/High NPI**, **Luminal B** subtype, Clusters **1, 6, 9** → RFS benefit
- **De-escalation considered**:
  - **Low/Intermediate NPI**, Cluster **9** → OS minimal benefit
  - **Low/Intermediate NPI**, Cluster **4** → minimal RFS benefit

Low NPI (<4.0)

- **Escalation recommended**:
  - **Young (<55)** or **Middle-aged (55–70)**, with immune-active **Cluster 4** → OS benefit
  - **Young (<55)** with proliferative/unstable Clusters **2, 6, 8, 9, 10** → RFS benefit
- **De-escalation considered**:
  - **Older (>70)**, genomically quiet **Clusters 8, 9** → OS limited benefit
  - **Older (>70)** with **Luminal A / Claudin-low**, Clusters **3, 4, 8** → limited RFS benefit