Validation of MRI predictive models for treatment optimization in the I-SPY 2 TRIAL

Wen Li PhD[1], Natsuko Onishi MD PhD[1], John Kornak PhD[1], Christina Yau PhD[1], Denise M. Wolf PhD[41] Jessica E. Gibbs BA[1], Teffany Joy Bareng BS[1], Nu N Le MS[1], Lisa J. Wilmes PhD[1], Pouya Metanat MD MPH[1], Matthew Gibbons PhD[1], Elissa Price MD[1], Bonnie N. Joe MD PhD[1], Sara Venters PhD[1], Keli Santos-Parker MD PhD[1], Barbara LeStage MHP[2], Mark J Magbanua PhD[1], Laura J. van 't Veer PhD[1], I-SPY2 Imaging Working Group, I-SPY2 Network, Angela M DeMichele MD[3], Laura J. Esserman MD MBA[1], Nola M. Hylton PhD[1]

[1]University of California, San Francisco, [2]I-SPY 2 Advocacy Group, [3]University of Pennsylvania

BACKGROUND

Functional tumor volume (FTV) calculated from dynamic contrast-enhanced MRI is an imaging biomarker for breast cancer response to neoadjuvant therapy (NAT) established in I-SPY trials[1,2].

The FTV-based MRI predictive model is part of the treatment redirection algorithm in I-SPY 2 and I-SPY 2.2 to proceed to surgery early to avoid unnecessary toxicity from their assigned drugs.

In this study, we present the validation of subtype-specific MRI predictive models in the I-SPY 2 TRIAL.

METHODS AND MATERIALS

Study cohort: patients enrolled and randomized to treatment in I-SPY 2 between May 2010 and June 2022.

NAT endpoint: assessed by surgical pathology, Residual Cancer Burden (RCB).

- Good response – RCB-0/I
- Poor response – RCB-II/III

Breast cancer subtype:

- Receptor subtype (4 categories): HR+/HER2-, HR+/HER2+, HR-/HER2+, and TNBC (triple negative breast cancer)
- Response-predictive subtype (RPS, 6 categories): S1 – HR+HER2-ImPrint-RePrint-, S2 – HR-HER2-ImPrint-RePrint-, S3 – HER2-ImPrint+, S4 – HER2-ImPrint-RePrint+, S5 – HER2+HER2orBasal, S6 – HER2+Luminal

MRI predictive model: Logistic regression with baseline FTV and FTV measurements at week 3 or week 12 (Figure 1) for predicting good vs poor responses. Variables were selected by subtype. Area under the ROC curve (AUC) was used to assess model performance.

Figure 1 MRI predictive model was built with FTV variables up to T2

RESULTS – Predicting Pathological Response

Receptor-subtype-specific models

Table 1 Comparison of model performance in the training and test set

Subtype	Training set n=1267			Test set n=556		
	N	Response rate	AUC (95% CI)	N	Response rate	AUC (95% CI)
HR+/HER2-	546	36% (197/546)	0.72 (0.68, 0.76)	225	36% (82/225)	0.66 (0.58, 0.73)
HR+/HER2+	192	60% (115/192)	0.72 (0.64, 0.79)	87	53% (46/87)	0.68 (0.56, 0.79)
HR-/HER2+	92	79% (73/92)	0.78 (0.67, 0.88)	43	84% (36/43)	0.65 (0.404, 0.89)
TNBC	437	62% (272/437)	0.77 (0.73, 0.82)	201	59% (119/201)	0.78 (0.72, 0.84)

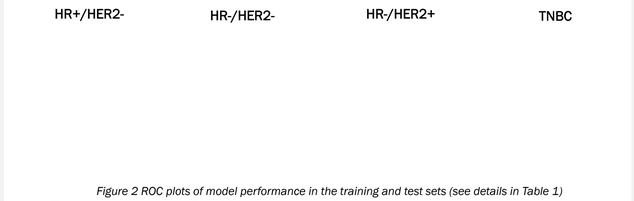


Figure 2 ROC plots of model performance in the training and test sets (see details in Table 1)

- The receptor-specific MRI models performed consistently in the test sets compared to the model performances in the training sets, across all subtypes.
- The largest difference in AUC (training vs test) was observed in HR-/HER2+ but it was not statistically significant (p=0.31).

Receptor- vs RPS-specific models

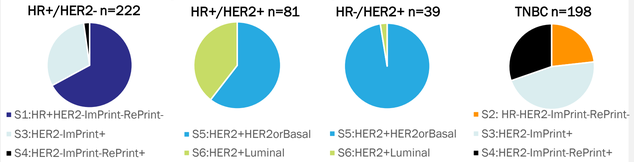


Figure 3 Comparison of receptor-specific vs RPS-specific models in the test set

RPS-subtype-specific models

Table 2 Comparison of model performance in the training and test set

Subtype	Training set n=1267			Test set n=556		
	N	Response rate	AUC (%95 CI)	N	Response rate	AUC (%95 CI)
S1: HR+HER2-ImPrint-RePrint-	370	24% (89/370)	0.72 (0.66, 0.78)	149	23% (35/149)	0.69 (0.58, 0.799)
S2: HR-HER2-ImPrint-RePrint-	97	40% (39/97)	0.81 (0.72, 0.89)	46	37% (17/46)	0.72 (0.56, 0.88)
S3: HER2-ImPrint+	332	72% (240/332)	0.72 (0.66, 0.78)	160	68% (108/160)	0.70 (0.61, 0.79)
S4: HER2-ImPrint-RePrint+	155	57% (88/155)	0.82 (0.75, 0.89)	65	58% (38/65)	0.85 (0.76, 0.95)
S5: HER2+HER2or Basal	202	76% (153/202)	0.74 (0.66, 0.73)	89	78% (69/89)	0.73 (0.74, 0.86)
S6: HER2+Luminal	70	37% (26/70)	0.72 (0.59, 0.85)	33	21% (7/33)	0.68 (0.48, 0.87)

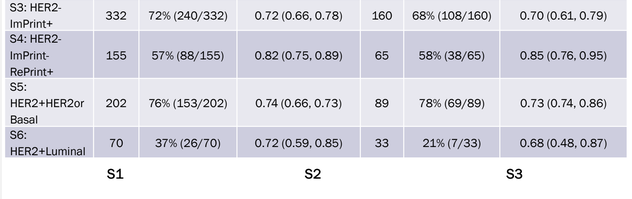


Figure 4 ROC plots of model performance in the training and test sets (see details in Table 2)

- Similar to receptor-specific MRI models shown on the left, the RPS-specific MRI models performed consistently in the test sets compared to the model performances in the training sets, across all subtypes.
- The largest difference in AUC (training vs test) was observed in S2: HR-HER2-ImPrint-RePrint-, but it was not statistically significant (p=0.34).

DISCUSSION

- This study validated that subtype-specific MRI models achieved consistent performance in the test cohort for predicting response.
- Compared to HR/HER2 subtypes, RPS is a refined classification of breast cancer based on molecular, genetic, or histological characteristics of the tumor that correlate with treatment outcomes.
- In our validation (test) cohort, higher AUCs were observed in HR+/HER2- and HR+/HER2+ sub-cohorts if RPS-specific models were used.
- The RPS-specific models are currently being used in combination with core biopsy as part of the clinical algorithm for tailoring therapy to response in I-SPY 2.2.
- **From patient advocate**: *Predicting patients' response by RPS can improve the effectiveness of MRI in measuring treatment response.* Early prediction of responders will increase treatment optimization by giving responders the option of omitting AC chemotherapy and going directly to surgery after their first treatment regimen. This allows patients to avoid the side and late effects of AC, improving their quality of life.

CONCLUSIONS

Subtype-specific MRI modeling can capture response heterogeneity, therefore enabling individualized breast cancer treatment. Improved MRI predictive performance was observed when HR/HER2 subtypes were refined by response-based RPS subtypes.

In ISPY2, MRI FTV is an important tool to improve prediction of response. We are working to standardize and disseminate tools for FTV assessment to facilitate adoption and validation.

REFERENCES:

1 Hylton et al. Radiology. 2012. 2 Li et al. Tomography. 2025.

ACKNOWLEDGEMENTS:

This research was supported by the NCI of the National Institutes of Health (NIH) under award number R01 CA132870, U01 CA225427, P01CA210961, and R01 CA255442. The authors wish to acknowledge the generous support of the 501(c)3 study sponsors, Quantum Leap Healthcare Collaborative and express appreciation for the ongoing support for the I-SPY program from the Safeway Foundation, the William K. Bowes, Jr. Foundation, Give Breast Cancer the Boot, and the Breast Cancer Research Foundation. Sincere thanks to our patient advocates and all the patients who volunteered to participate in the I-SPY trials.

This presentation is the intellectual property of the author/presenter. Contact them at wen.li@ucsf.edu for permission to reprint and/or distribute.