Advancing Breast Cancer Research in Japan Using EHR-Derived Real-World Data

Blythe Adamson¹, Dionne Ng², Harlan Pittell¹, Arun Sujenthiran³, Eri Tajima²

¹Flatiron Health, New York, NY; ²Flatiron Health K.K., Tokyo, Japan; ³Flatiron Health UK, London, UK

Background

- Breast cancer is the most commonly diagnosed cancer and the fourth leading cause of cancer-related death among women in Japan, with significant heterogeneity in genomic subtypes, treatment patterns, and clinical outcomes²
- Real-world evidence (RWE) is critical for understanding treatments and outcomes in routine clinical practice³
- This study aimed to characterize a growing cohort as a comprehensive, deidentified resource for evaluating clinical profiles and treatment effectiveness among patients with breast cancer in Japan receiving care in real-world settings

Methods

- **Data source:** The Flatiron Health Research Database, an electronic health record (EHR)-derived, deidentified database, comprising patient-level data originating from over 5 million patients around the world, including the US, Germany, UK, and Japan⁴⁻⁵
 - The Japanese dataset was used for this study. New patients were added every 90 days, and new information for existing patients was updated with 90-day recency, based on technology-enabled extraction and abstraction of relevant information from both structured and unstructured EHR⁵ data sources
- Eligibility criteria: The study included 1057 patients in Japan with a confirmed diagnosis of breast cancer (ie, abstraction-confirmed pathology, stage, etc.) between January 1, 2011, and September 30, 2024
- Analyses: We summarized the characteristics of this Japanese cohort using descriptive statistics. Variables examined included:
 - Clinical characteristics: Age at diagnosis, birth sex, tumor laterality, clinical or pathological group stage, tumor grade, histology, metastatic status, and Eastern Cooperative Oncology Group performance status (ECOG PS) score at diagnosis
 - Interventions: Biomarker testing rates and results, surgery, and radiation therapy

Results

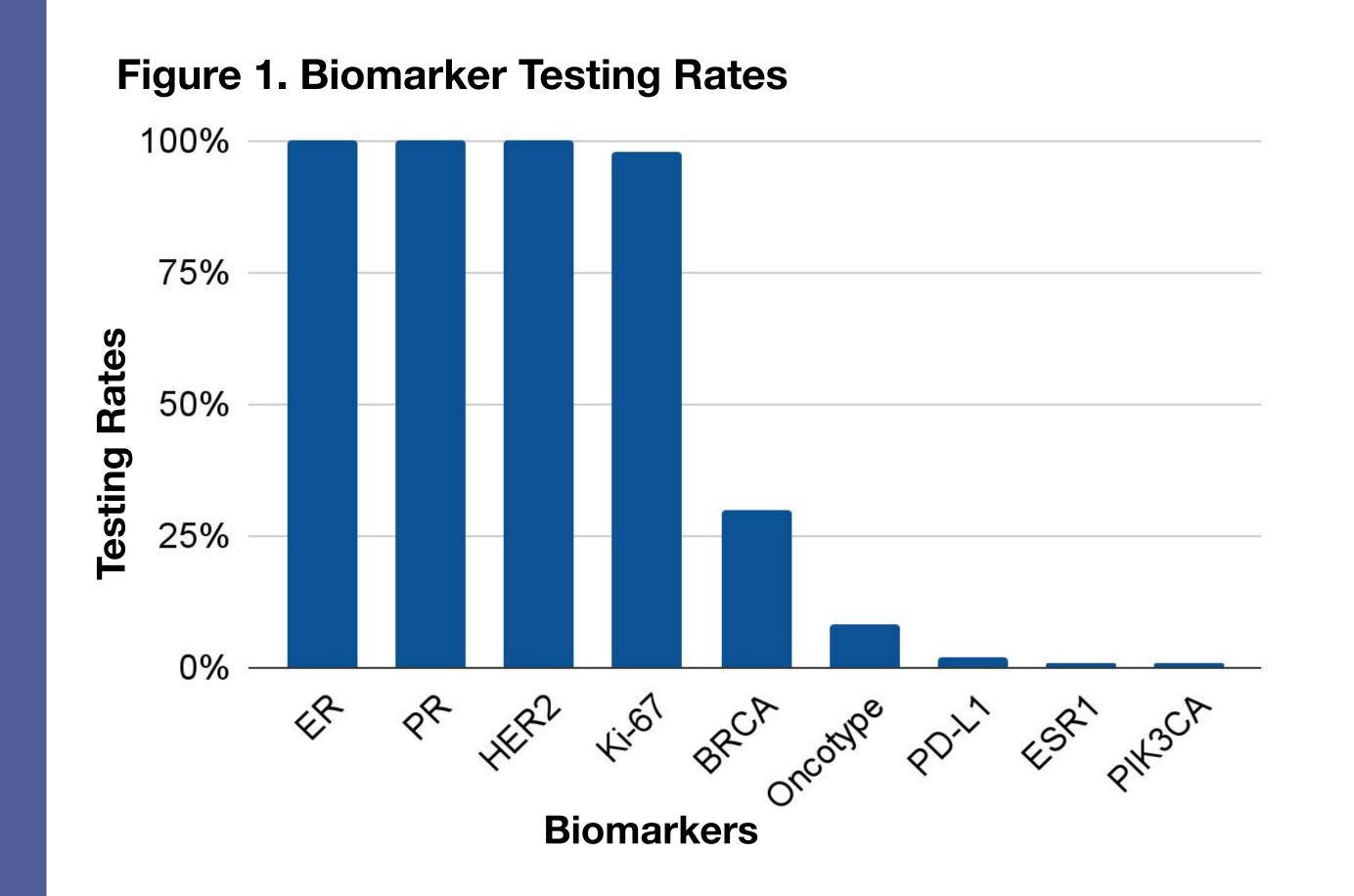
Table 1. Patient Characteristics

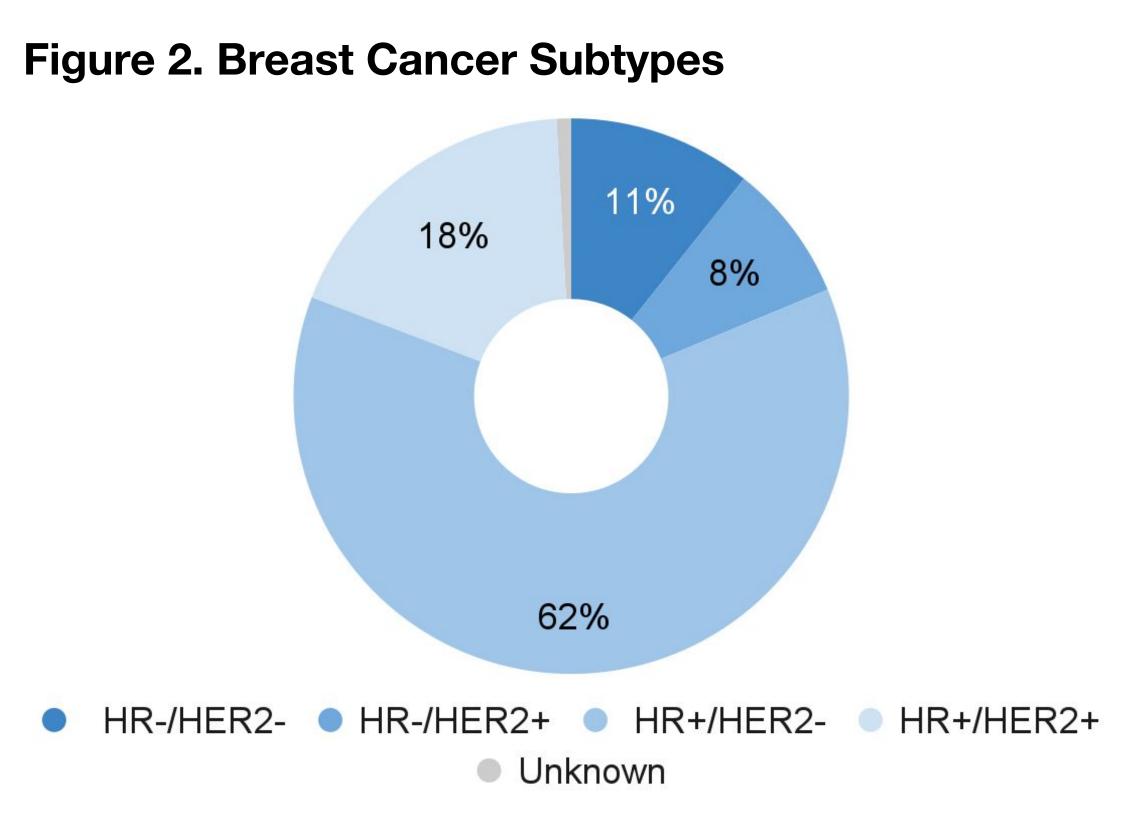
Characteristic	N = 1057
Age , median (IQR), y	54 (46-65)
Sex, %	
Female	100
Tumor laterality, No. (%)	
Left	548 (52)
Right	509 (48)
Menopausal status, No. (%)	
Premenopausal	415 (39)
Postmenopausal	414 (39)
Other/Unknown ^a *	228 (22)
Group stage ^b , No. (%)	
Stage 0-I	253 (24)
Stage II	477 (45)
Stage III	178 (17)
Stage IV	96 (9)
Unknown	53 (5)

Characteristic	N = 1057
Tumor grade, No. (%)	
Grade 1	146 (14)
Grade 2	420 (40)
Grade 3	175 (17)
Unknown	316 (30)
Histology, No. (%)	
Invasive/Infiltrating ductal carcinoma	923 (87)
Invasive/Infiltrating lobular carcinoma	69 (7)
Other	54 (5)
Unknown	11 (1)
Metastatic status, No. (%)	212 (20)
De novo metastatic	97 (9)
Distant recurrence	115 (11)
Availability of metastatic diagnosis date, No. (%)	212 (100) ^c
ECOG PS 0 or 1 at first treatment, %	~99

Results (continued)

- The majority of patients were tested for *ER*, *PR*, *HER2*, and *Ki-67*, whereas other standard of care biomarkers were tested less frequently in this cohort
- Overall, 26% of patients had HER2+ disease, 56% had HER2-low disease, and 11% had triple-negative breast cancer
- A total of 87% of patients underwent surgery, and 52% received radiation therapy





Abbreviations: *BRCA*, breast cancer gene; ER, estrogen receptor; *ESR1*, estrogen receptor 1; HER2, human epidermal growth factor receptor 2; *HER2*+, human epidermal growth factor receptor 2 negative; *HR*+, hormone receptor positive; *HR*-, hormone receptor negative; *Ki-67*, proliferation index marker Ki-67; PD-L1, programmed death-ligand 1; *PIK3CA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; PR, progesterone receptor

Conclusion and Future Directions

- This is the first known study to examine detailed clinical and biomarker characteristics in a real-world cohort of patients diagnosed with breast cancer in Japan, using an EHR-derived real-world dataset
- The database contained detailed clinical variables with high completeness, and key characteristics of this Japanese breast cancer cohort were largely consistent with clinical expectations^{6,7}
- Secure analyses of EHR-derived, deidentified, patient-level Japanese RWD in a trusted research environment enable robust evidence generation, in compliance with local legal and ethical requirements, while protecting patient privacy
- Future research will focus on comparative effectiveness and cross-country analyses to advance global breast cancer research

References

- 1. Ganjoho. 最新がん統計. ganjoho.jp. Updated December 16, 2024. Accessed April 25, 2025. https://ganjoho.jp/reg_stat/statistics/stat/summary.html
- 2. Nagahashi M, et al. *Breast Cancer*. 2025;32(2):217-226. doi:10.1007/s12282-025-01671-0
- 3. Bando H, et al. *ESMO Real World Data Digit Oncol*. 2023;2:100005. doi:10.1016/j.esmorw.2023.100005
- 4. Flatiron Health. Database Characterization Guide. Flatiron.com. Published March 18, 2025. Accessed April 16, 2025. https://flatiron.com/database-characterization
- 5. Adamson B, et al. ESMO Real-World Data and Digital Oncology. 2025. doi:10.1016/j.esmorw.2025.100113
- 6. Japan Breast Cancer Society. BQ4 What is the purpose of HER2 testing and how is it performed? jbcs.xsrv.jp. Published November 15, 2022. Updated March 28, 2024. Accessed April 25, 2025. https://jbcs.xsrv.jp/guideline/2022
- 7. Abdou Y, et al. *Sci Rep.* 2022;12(1):13377. doi:10.1038/s41598-022-16749-4

Acknowledgments: Darren Johnson for project management and publication support

Disclosures: This study was sponsored by Flatiron Health, Inc.—an independent member of the Roche Group. During the study period, BA, DN, HP, AS, and ET reported employment with Flatiron Health, Inc., and stock ownership in Roche. Data first presented at ISPOR 2025 in Montreal, QC, Canada on May 16, 2025

Author contact information: Blythe Adamson badamson@flatiron.com



Scan to learn

Abbreviations: IQR, interquartile range

^aOther/Unknown includes patients with perimenopausal or N/A (patient is male) or Unknown status, with the former 2 categories aggregated due to small patient counts that cannot be reported for privacy reasons.

^bAggregated group stage based on clinical and pathological stage data.

^cDenominator = patients with metastatic disease.