AACR
American Association for Cancer Research®
FINDING CURES TOGETHER®

PROJECT GENIE®
Genomics Evidence Neoplasia Information Exchange

GENIE BPC BRCA1.0-PUBLIC COHORT

November 2025

AACR.org/GENIE • #AACRGENIE

Release Notes

- The GENIE BPC BRCA 1.0-public dataset contains 1,130 breast cancer patients from 3 institutions: DFCI, MSK, and VICC
- **Data Access:**
  - A subset of the data is available through cBioPortal. Both the genomic and clinical (phenomic) data can be evaluated in cBioPortal with opportunities for data exploration and visualization using a user-friendly interface.
  - The complete, post-processed data are available on Synapse.

- **What is included in GENIE BPC data?**
  - **Genomic Data:** Clinical-grade next-generation sequencing data for each patient from the GENIE Registry. Genomic profiling was performed between 2013 and 2018; patients were aged 18-56 at the time of genomic sequencing and include male breast cancers.
  - **Cancer Diagnosis:** Breast cancer diagnosis is considered the index tumor for this patient cohort. There are data about other cancer diagnoses antecedent to the breast cancer and subsequent to the breast cancer.
    - Breast cancer specific fields such as, ER (Estrogen Receptor) status, PR (Progesterone Receptor) status, HER2 (Human Epidermal Growth Factor Receptor 2) status and summary, Oncotype DX recurrence score, multigene signature method and results, band breast subtype classification were also collected.
  - **Pathologic Information:** Each pathology specimen from diagnosis through death or last follow-up is curated with specimen type, site, and histology.

Release Notes

- **Treatment Histories:** All anti-neoplastic systemic therapies–intravenous and oral chemotherapies–are included in the data set. Dates are provided as intervals from diagnosis to start and stop of each drug. Investigational drugs are masked, no dosing information is included.
- **Imaging Information:** Each CT, MRI, PET-CT scan from diagnosis through death or last follow-up is curated for the presence or absence of cancer and an evaluation of whether the cancer was stable, responding, or progressing. These data are used to compute progression-free survival-imaging (PFS-I). Sites of tumor involvement are also recorded.
- **Medical Oncologist's Evaluations:** Medical oncology notes (1/month) have been curated to ascertain the presence or absence of cancer and whether the cancer was stable, responding, or progressing. These data are used to compute progression-free survival-medonc (PFS-M) from diagnosis through death or date of last follow-up. The ECOG or Karnofsky Performance Status (KPS) were curated when available in the medical oncology note.
- **Overall Survival:** Overall survival is based on death, with censoring at the date last known alive. Ascertainment of death varies by institution.
- **Additional Relevant Biomarkers:** Information about select biomarkers not included on the NGS panels, including PD-L1, MMR, and MSI are also curated.
- **Patient-Reported Outcomes:** No patient-reported outcomes are available in this dataset.
- **Date Masking:** Exact dates are masked to preserve confidentiality; however, date intervals are available, allowing calculation of event times such as diagnosis, treatment start, treatment end, PFS-I, PFS-M, and OS

Release Notes

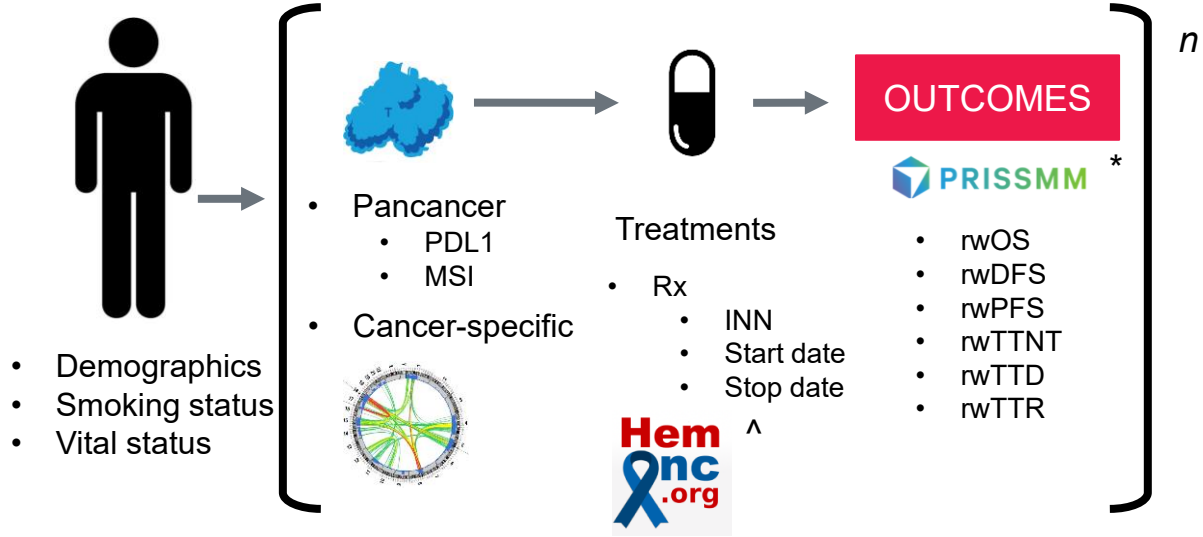
- **Analytical Data Guide:** A more comprehensive overview of the data can be found in the data guide, and a description and location of the variables collected can be found in the variable synopsis spreadsheet.

- **Other Resources:** There is a dedicated project wiki that describes each of the files.

- **Training Videos**:
  - Demo of GENIE Data on the Synapse and cBioPortal Platforms: here
  - BPC- specific cBioPortal video training playlist: here

- **PRISSMM™:** the BPC BrCa dataset uses the PRISSMM™ system licensed and enhanced by Memorial Sloan-Kettering Cancer Center, Memorial Hospital for Cancer and Allied Diseases, and Sloan-Kettering Institute for Cancer Research (collectively "MSK") is for informational and research purposes only. The content is not intended as a substitute for professional medical advice, diagnosis, or treatment. Original system and improvements © 2019-2022 Dana-Farber Cancer Institute, Inc. Additional functionality and enhancements © 2023 MSK. All rights reserved. Additional information can be found in the analytic data guide and information about licensing PRISSMM™ can be obtained by emailing PRISSMM@mskcc.org

GENIE Data Model



*PRISSMM is licensed from the DFCI
^HemOnc is available through a creative commons license

GENIE BPC Data Model

Core Patient-Level Data
- Demographics
- Smoking history
- Vital status

n

Biomarkers

Pan-cancer biomarkers
- PD-L1
- MSI

Cancer-specific attributes
- Histology
- Stage
- Receptor status

Treatment Data

Systemic Therapies (Rx)
- International Nonproprietary Name (INN)
- Start and stop dates

Lines of therapy
- Standardized treatment ontology
- HemOnc.org (Creative Commons license)

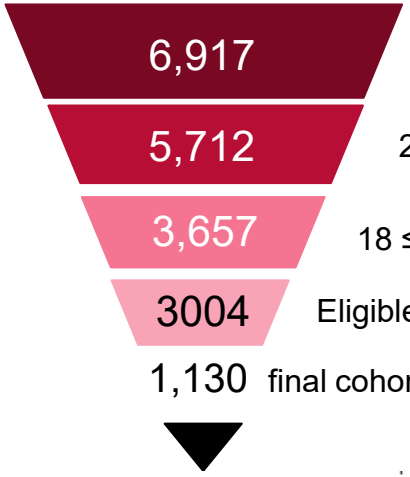
Clinical Outcomes

PRISSMM

rwOS – Real-world Overall Survival
rwDFS – Disease-Free Survival
rwPFS – Progression-Free Survival
rwTTNT – Time to Next Treatment
rwTTD – Time to Treatment Discontinuation
rwTTR – Time to Response

PRISSMM™ is licensed from the Dana-Farber Cancer Institute.

BPC BrCa 1.0-public Cohort Preview



Breast Diagnoses at BPC Phase 1 Institutions

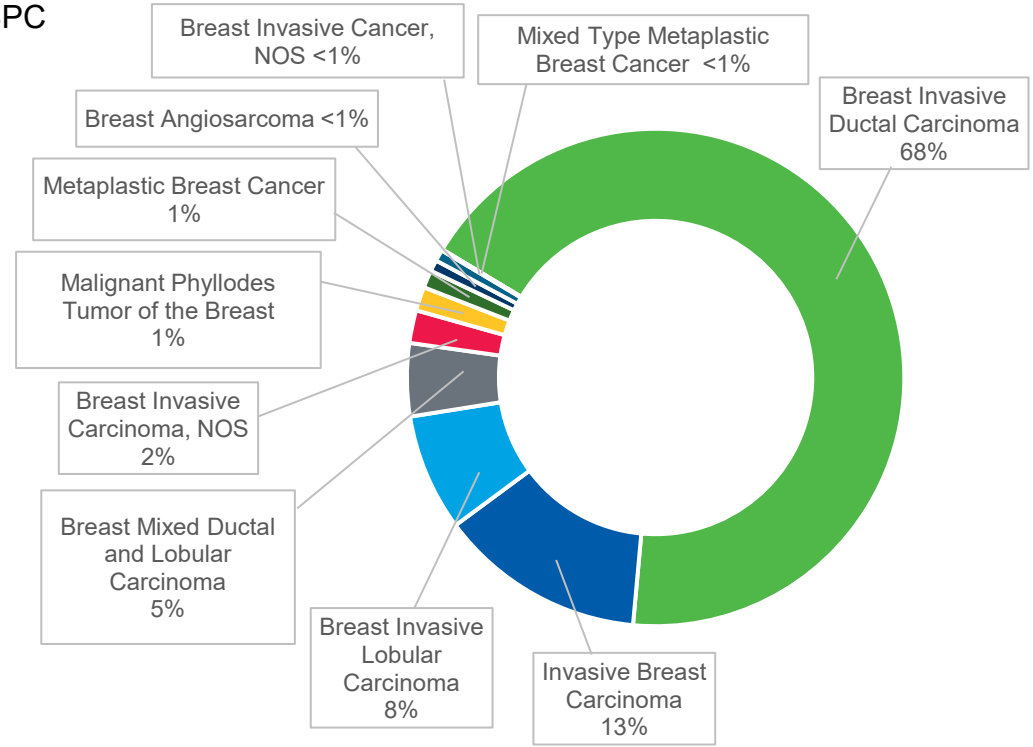
2013 ≤ NGS ≤2018

18 ≤ Age ≤ 56

Eligible & randomly selected

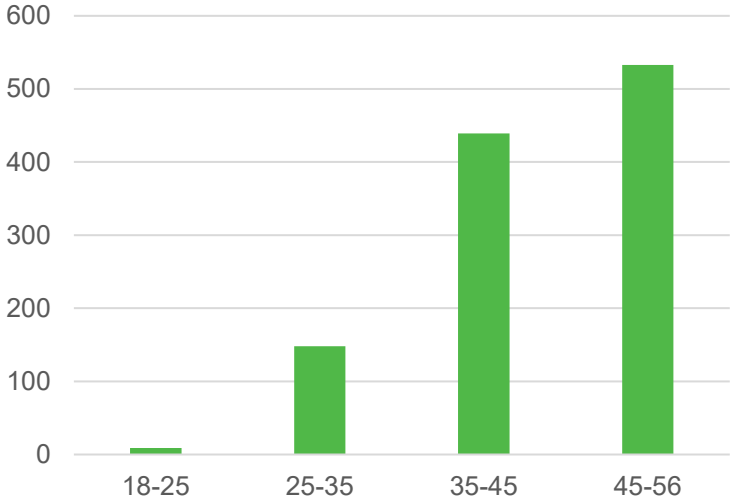
1,130 final cohort (1,223 samples total)

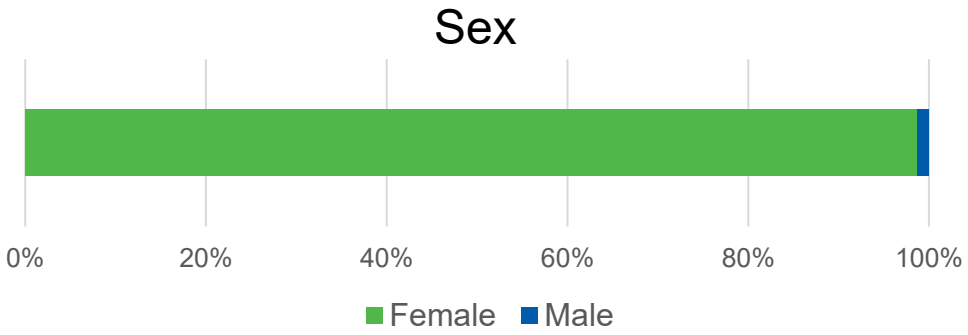
Breast Invasive Ductal Carcinoma	830	67.9%
Invasive Breast Carcinoma	164	13.4%
Breast Invasive Lobular Carcinoma	93	7.6%
Breast Mixed Ductal and Lobular ...	58	4.7%
Breast Invasive Carcinoma, NOS	26	2.1%
Malignant Phyllodes Tumor of the...	19	1.6%
Metaplastic Breast Cancer	14	1.1%
Breast Angiosarcoma	9	0.7%
Breast Invasive Cancer, NOS	9	0.7%
Mixed Type Metaplastic Breast C...	1	<0.1%

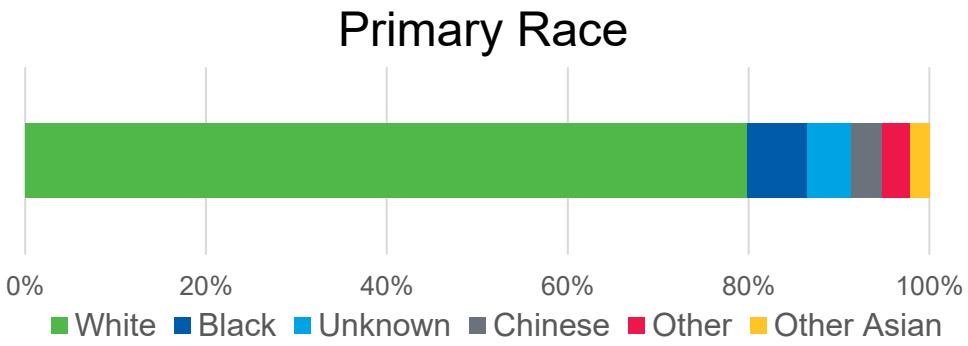


**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public Demographics

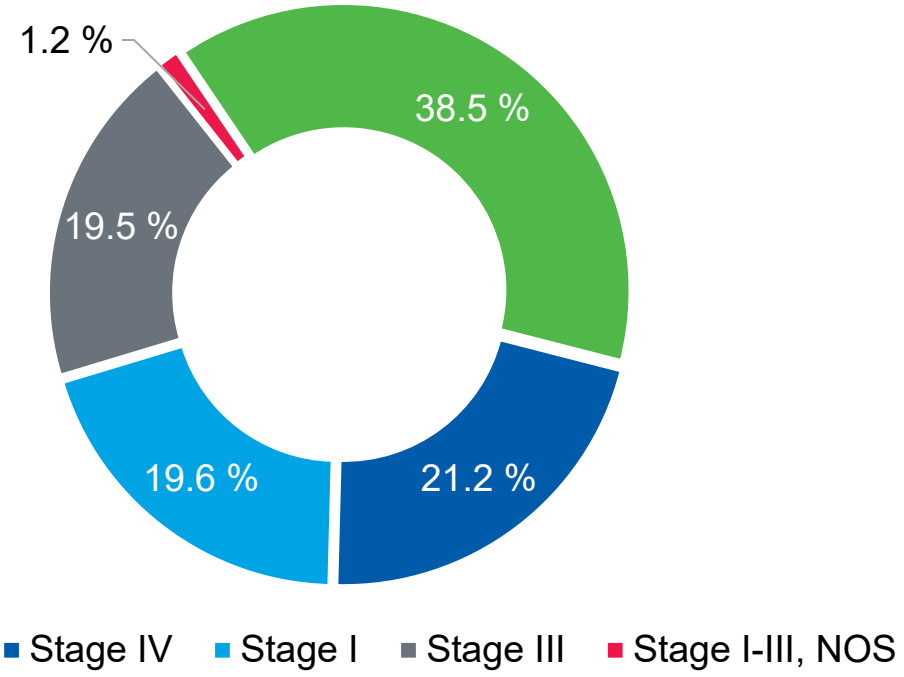






**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

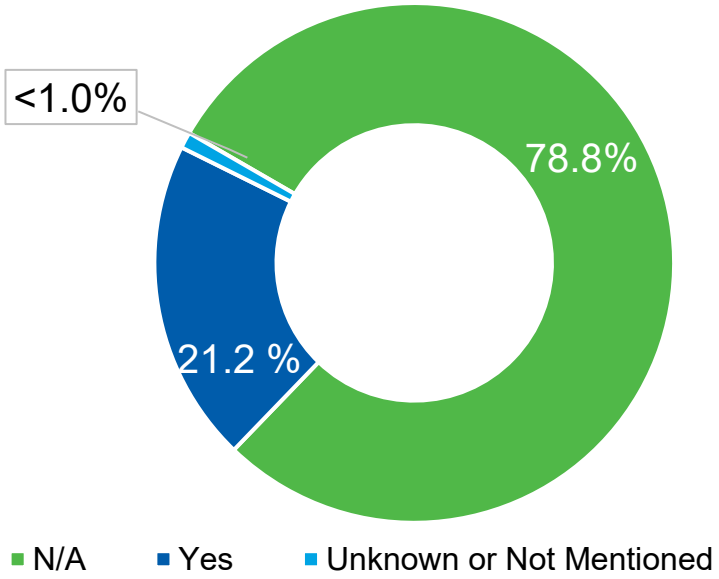
BPC BrCa 1.0-public: Stage at Diagnosis



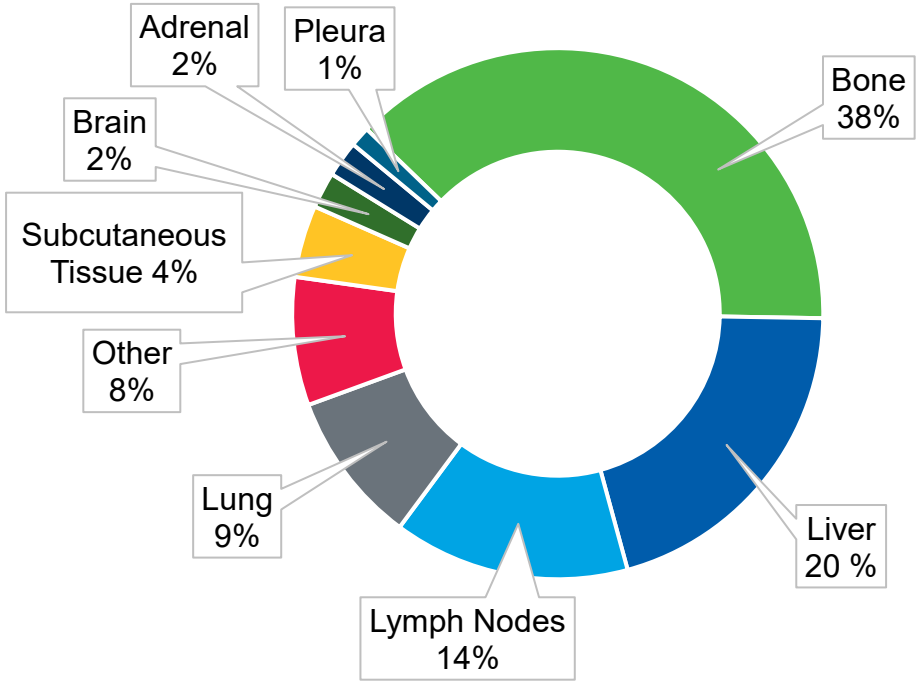
*BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®

BPC BrCa 1.0-public: Sites of Metastases at Diagnosis

Patients with Distant Metastases at Diagnosis



Distribution of Sites of Distant Metastasis



**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: Detailed Clinical Genomics

AACR American Association for Cancer Research FINDING CURES TOGETHER | PROJECT GENIE Genomics Evidence Neoplasia Information Exchange

Mutated Genes (1223 profiled samples)

Gene	# Mut	#	Freq
TP53	584	565	46.2%
PIK3CA	397	350	28.6%
GATA3	187	176	14.4%
CDH1	125	122	10.0%
ARID1A	90	84	6.9%
KMT2D	84	77	6.3%
MAP3K1	91	74	6.1%
PTEN	86	73	6.0%
BRCA2	72	68	5.6%
ESR1	66	63	5.2%
NF1	65	62	5.1%

Structural Variant Genes (1221 profiled samples)

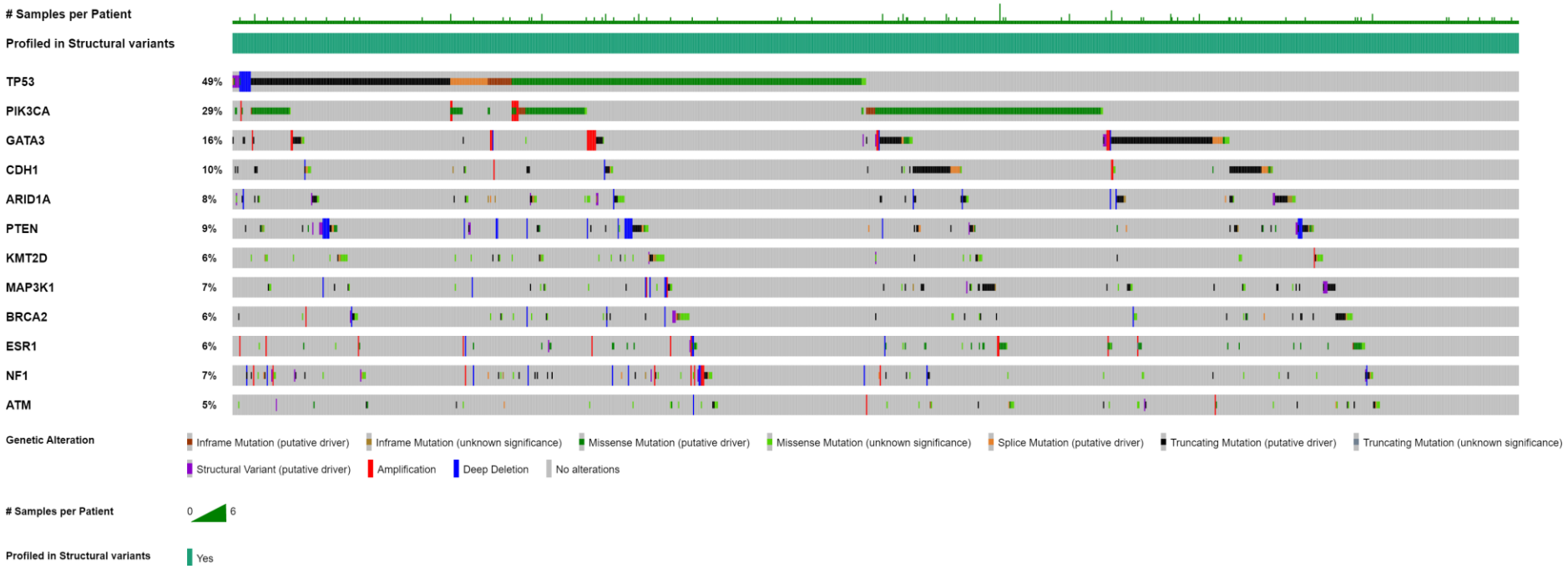
Gene	# SV	#	Freq
CDK12	11	11	1.2%
PTEN	10	10	0.8%
ERBB2	10	10	0.8%
BRIP1	8	8	0.7%
NOTCH2	8	8	0.7%
FGFR2	11	8	0.7%
FGFR1	8	8	0.7%
RARA	7	7	0.6%
ARID1A	7	7	0.6%
RB1	6	6	0.5%
MYC	6	6	0.5%

CNA Genes (1223 profiled samples)

Gene	Cytoband	CNA	#	Freq
CCND1	11q13.3	AMP	209	17.1%
ERBB2	17q12	AMP	180	14.7%
MYC	8q24.21	AMP	153	12.5%
FGFR1	8p11.23	AMP	145	11.9%
FGF19	11q13.3	AMP	143	18.1%
FGF4	11q13.3	AMP	137	17.3%
FGF3	11q13.3	AMP	135	17.1%
CDK12	17q12	AMP	107	11.3%
RAD21	8q24.11	AMP	83	8.5%
NSD3	8p11.23	AMP	61	11.3%
BRIP1	17q23.2	AMP	59	4.8%

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: Top 12 Mutated Genes



*BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®

BPC BrCa 1.0-public: Top Clinically Actionable Genes

Click here for more information on: Actionable Genes

OncoKB™

Chakravarty et al., JCO PO 2017

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: HER2 Receptor Status Summary

HER2 Summary Result: Overall human epidermal growth factor receptor 2 (HER2) status at diagnosis. This variable accounts for both IHC and ISH results at diagnosis.

HER2 IHC Lab Value: Breast cancer human epidermal growth factor receptor 2 (HER2) immunohistochemistry lab value at diagnosis (Collaborative Stage Site Specific Factor 8)

HER2 IHC Lab Interpretation at Diagnosis: Breast cancer human epidermal growth factor receptor 2 (HER2) immunohistochemistry lab interpretation at diagnosis (Collaborative Stage Site Specific Factor 9)

HER2 Summary Result

	#	Freq
Negative/normal within normal limits	849	75.1%
Positive/elevated/amplified	215	19.0%
Unknown or no information	34	3.0%
Borderline/equivocal/indeterminant	14	1.2%
NA	10	0.9%
Tests not done	7	0.6%
Test ordered results not in chart	1	<0.1%

HER2 IHC Lab Value

	#	Freq
Score 0	302	26.7%
Score of 1+	301	26.6%
Unknown or no information	171	15.1%
Score of 3+	143	12.7%
Score of 2+	125	11.1%
Test not done	74	6.5%
NA	10	0.9%
Test ordered/results not in chart	4	0.4%

HER2 IHC Lab Interpretation at Diagnosis

	#	Freq
Negative/normal	692	61.2%
Positive/elevated	156	13.8%
Borderline; equivocal; indetermin...	131	11.6%
Test not done	73	6.5%
Unknown or no information	64	5.7%
NA	12	1.1%
Test ordered/results not in chart	2	0.2%

BPC BrCa 1.0-public: Hormone Receptor Status Summary

ER Summary Status: Breast cancer estrogen receptor (ER) summary (Collaborative Stage Site Specific Factor 1)

PR Summary Status: Breast cancer progesterone receptor (PR) summary (Collaborative Stage Site Specific Factor 2)

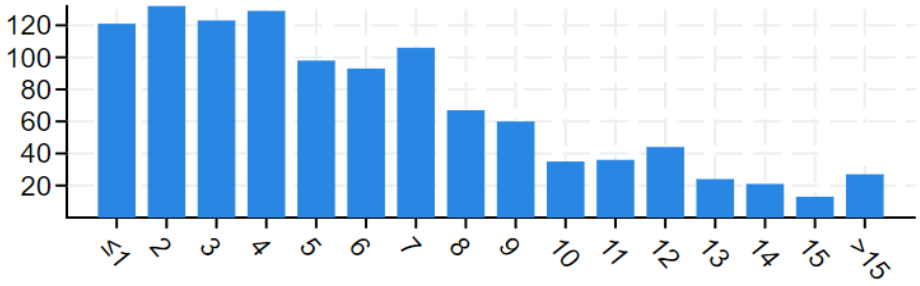
ER Summary Status	#	Freq
Positive/elevated	804	71.2%
Negative/normal	293	25.9%
Unknown or no information	17	1.5%
NA	9	0.8%
Test not done	4	0.4%
Test ordered/results not interpreta…	2	0.2%
Borderline	1	<0.1%

PR Summary Status	#	Freq
Positive/elevated	722	63.9%
Negative/normal	370	32.7%
Unknown or no information	18	1.6%
NA	9	0.8%
Test not done	7	0.6%
Borderline	3	0.3%
Test ordered/results not in chart	1	<0.1%

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public : Complete Treatment Histories

Number of Cancer-Directed Drug Regimens Curated for Each Patient



Treatment by Sample (pre- and post-NGS): useful to identify treatment induced mutations

Treatment	Pre / Post	#
Capecitabine	Pre	435
Capecitabine	Post	99
Cyclophosphamide	Pre	355
Cyclophosphamide	Post	428
Paclitaxel	Pre	402
Paclitaxel	Post	355
Doxorubicin HCL	Pre	318
Doxorubicin HCL	Post	381
Tamoxifen	Pre	355
Tamoxifen	Post	293
Investigational Drug	Pre	350

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: Complete Treatment Histories

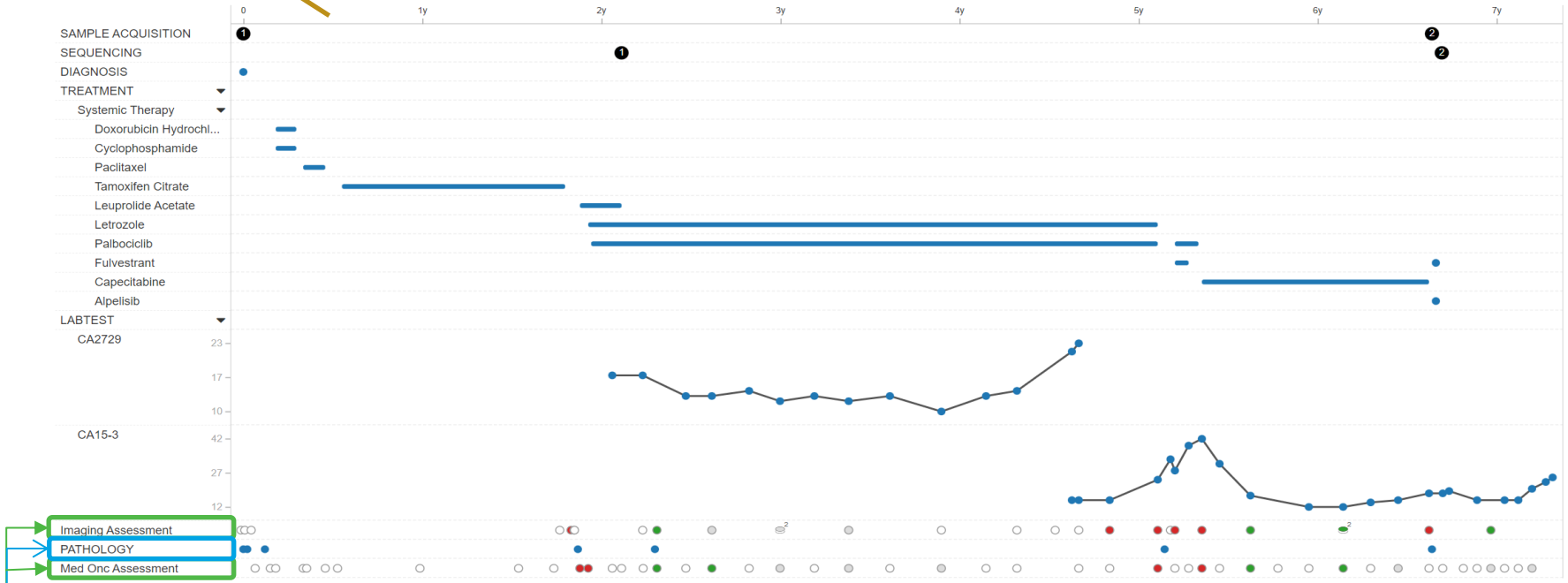
Cancer-directed Drug Regimens*	# of patients
Dose-Dense AC-T Chemotherapy (doxorubicin, cyclophosphamide, and paclitaxel)	532
TH Chemotherapy and Targeted Therapy (paclitaxel and trastuzumab)	191
TC Chemotherapy (docetaxel and cyclophosphamide)	171
CMF Chemotherapy (cyclophosphamide, methotrexate, and fluorouracil)	80

HER2-Directed Treatment Histories*	# of patients
T-DM1: Trastuzumab-DM1 (Trastuzumab emtansine)	101
Pertuzumab & T-DM1: Pertuzumab & Trastuzumab-DM1 (Trastuzumab emtansine)	81
L+T: Lapatinib & Trastuzumab	69
Neratinib monotherapy	12

*(includes regimens ever received for Breast cancer diagnosis irrespective of line or stage)

BPC BrCa 1.0-public: Comprehensive Patient View

Sample Patient

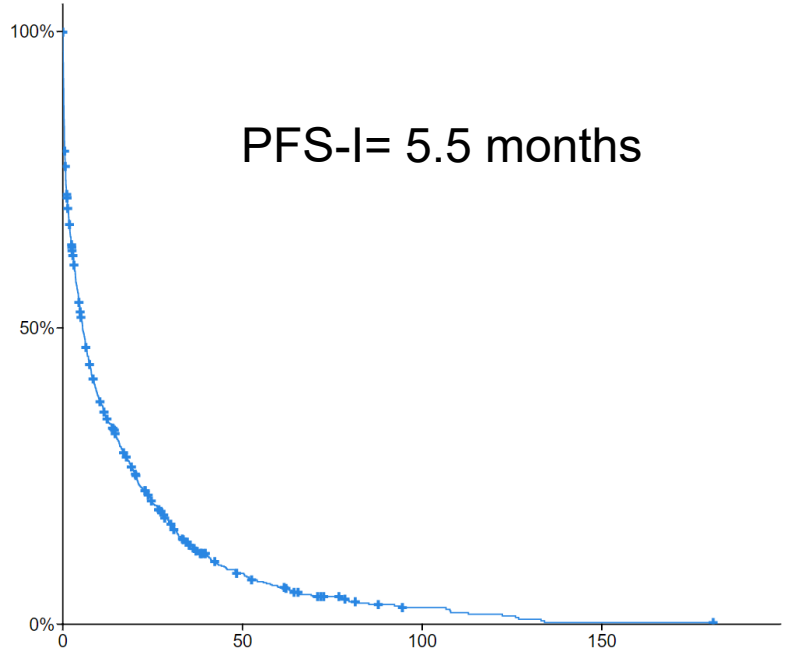


Pathology Reports

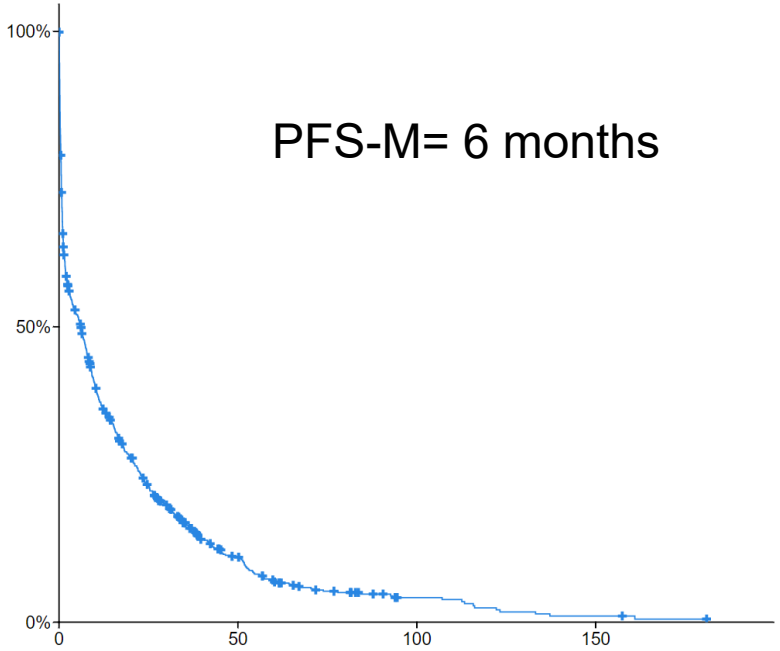
Disease Status by Medical Oncology Note and by Imaging

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: High Quality Clinical Data



PFS-I= 5.5 months



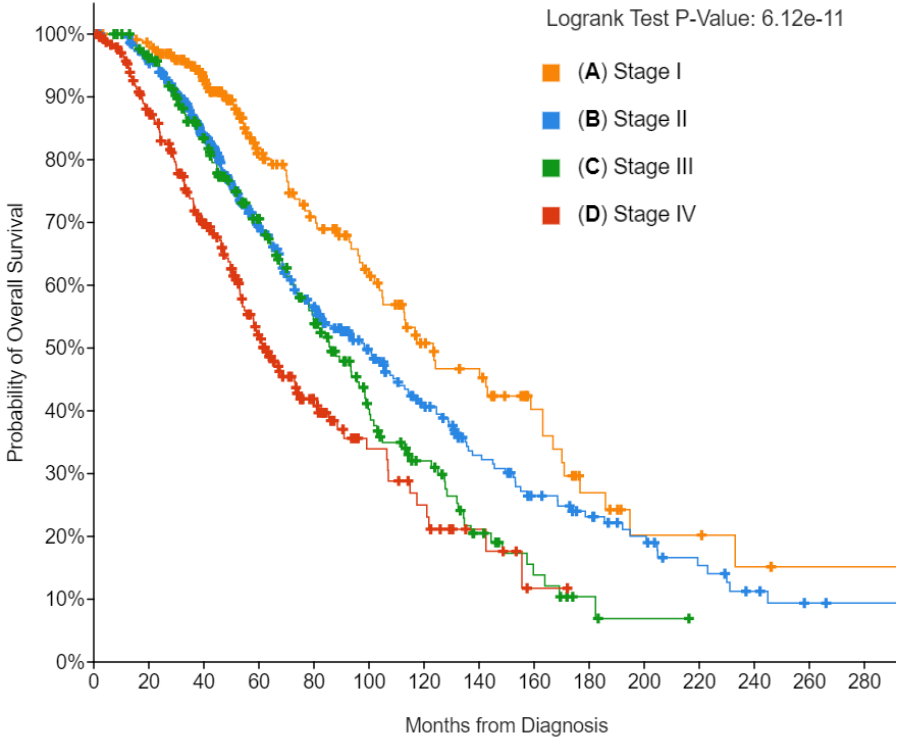
PFS-M= 6 months

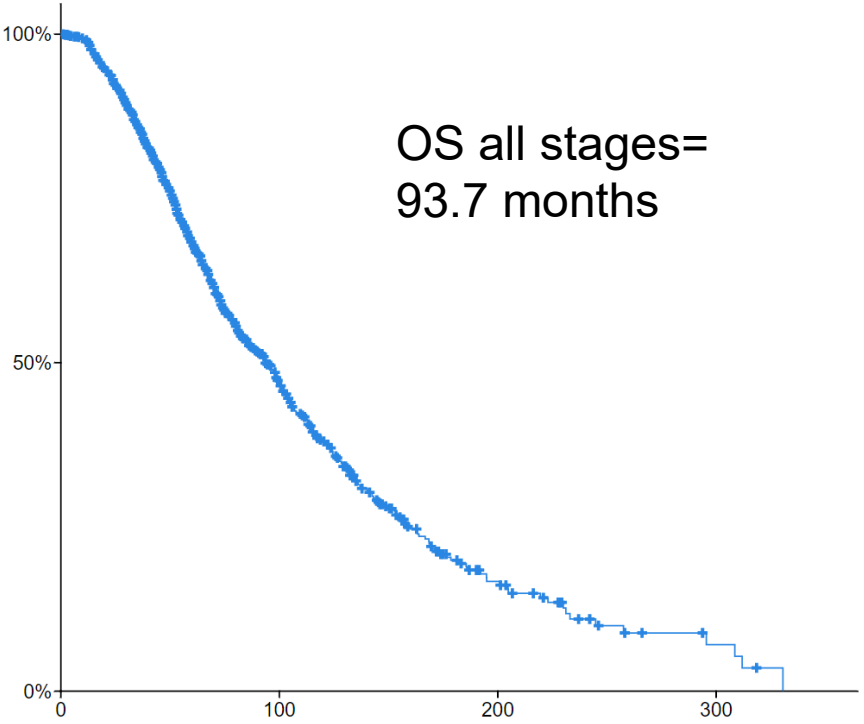
- PFS-M and PSF-I are available by regimen for those regimens containing greater than 10 patients as well as the entire cohort

* Not adjusted for delayed entry

**BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®*

BPC BrCa 1.0-public: Overall Survival by Stage





OS all stages=
93.7 months

BrCa 1.0-public Cohort ©2025 American Association for Cancer Research Project GENIE ®

*not adjusted for delayed entry

GENIE BPC Acknowledgements

Shawn M. Sweeney
Kelli Rasmussen
Alyssa Acebedo
Jennifer Hoppe

Niki Schultz
Ben Gross
Ritika Kundra
Brooke Mastrogiacomo

Xindi Guo
Thomas Yu
Chelsea Nayan
Alex Paynter

Gregory Riely
Deb Schrag
Julia Rudolph
Charles L. Sawyers
Hira Rizvi
John Phillip
Julian Schwartz
Marufur Bhuiya
Stu Gardos
Cynthia Chu
Shirin Pillai

Statistical Core
Kathy Panageas
Jessica Lavery
Samantha Brown
Hannah Fuchs
Axel Martin
Michael Curry

Ken Kehl
Asha Postle
Kevin Haigis
John Orechia
Daniel Quinn
Simon Arango Baquero

Philippe Bedard
Celeste Yu
Samanta Del Rossi
Nitthusha Singaravelan
Demi Plagianakos
Alisa Nguyen
Nazish Qazi
Gunjan Srivastava
Sophie Cooke
Alisha Rizvi

Christine Micheel
Rhonda Potter
Ben Ho Park
Sanjay Mishra
Daniel Fabbri
Marilyn Holt
Neha Jain
Protiva Rahman
Li Wen
Yuanchu James Yang
Kate Mittendorf

Former BPC Team Members can be found here