Sister Study Breast Cancer Events (any type) [1] Stratified by Race/Ethnicity [2]: Data Release 9.0 [3]

Characteristic	Non-Hispanic White N (total = 3,555)	%	Black (includes Hispanic Black) N (total = 337)	%	Non-Black Hispanic/Other N (total = 257)	%
Breast Cancer Subtype for 1st Diagnosis						
Invasive [4]	2,666	75.0	241	71.5	200	77.8
Ductal only (no lobular component) [5]	1,888	53.1	156	46.3	140	54.5
Lobular only (no ductal component) [5]	247	6.9	9	2.7	14	5.4
Ductal and lobular [5,6]	190	5.3	8	2.4	13	5.1
Inflammatory	5	0.1	1	0.3	0	0.0
Paget's	8	0.2	2	0.6	0	0.0
Phyllodes [7]	5	0.1	0	0.0	0	0.0
Other/unknown if ductal or lobular [8]	323	9.1	65	19.3	33	12.8
In situ	850	23.9	89	26.4	54	21.0
Ductal only (no lobular component) [9]	702	19.7	79	23.4	43	16.7
Lobular only (no ductal component) [9]	112	3.2	6	1.8	8	3.1
Ductal and lobular [6,9]	18	0.5	0	0.0	0	0.0
Paget's	5	0.1	0	0.0	0	0.0
Other/unknown if ductal or lobular [10]	13	0.4	4	1.2	3	1.2
Invasiveness unknown [11]	39	1.1	7	2.1	3	1.2
Stage at 1st Breast Cancer Diagnosis [12]						
0	769	24.1	74	28.2	46	20.7
I	1,644	51.5	116	44.3	113	50.9
II	610	19.1	60	22.9	43	19.4
III	150	4.7	12	4.6	18	8.1
IV	21	0.7	0	0.0	2	0.9
Unknown, missing, cannot stage	*361*		*75*		*35*	
Estrogen Receptor (ER) Status for 1st Breast Cancer Diagnosis [12]						
Positive	2,574	85.5	182	75.2	167	83.1
Negative	431	14.3	58	24.0	34	16.9
Borderline	5	0.2	2	0.8	0	0.0
Unknown, missing, not performed	*545*		*95*		*56*	
Progesterone Receptor (PR) Status for 1st Breast Cancer Diagnosis [12]						
Positive	2,137	73.8	138	59.7	133	67.9
Negative	751	25.9	92	39.8	63	32.1
Borderline	7	0.2	1	0.4	0	0.0
Unknown, missing, not performed	*660*		*106*		*61*	

Characteristic	Non-Hispanic White N (total = 3,555)	Non-Hispanic White %	Black (includes Hispanic Black) N (total = 337)	Black (includes Hispanic Black) %	Non-Black Hispanic/Other N (total = 257)	Non-Black Hispanic/Other %
HER2 Amplification Status for 1st Breast Cancer Diagnosis [12]						
Positive	321	12.9	30	15.6	34	19.2
Negative	2,137	85.5	160	83.3	141	79.7
Borderline	40	1.6	2	1.0	2	1.1
Unknown, missing, not performed	*1,057*		*145*		*80*	
Combined ER, PR, HER2 Status [12]						
Triple negative (ER-, PR-, and HER2-)	235	8.0	40	16.8	19	9.7
Not triple negative	2,689	92.0	198	83.2	177	90.3
Unknown, missing, not performed	*631*		*99*		*61*	
Diagnosis Confirmation [13]						
Medical	2,874	80.8	224	66.5	194	75.5
NDI Plus or death certificate	31	0.9	12	3.6	2	0.8
None	650	18.3	101	30.0	61	23.7
Age (in years) at 1st Breast Cancer Diagnosis [14]						
Under 50	338	9.5	49	14.5	35	13.6
50-64	1,596	44.9	160	47.5	148	57.6
65 and older	1,621	45.6	128	38.0	74	28.8
Menopausal Status at 1st Breast Cancer Diagnosis						
Pre-menopausal	546	15.4	61	18.2	56	21.8
Post-menopausal	2,993	84.6	274	81.8	201	78.2
Unknown	*16*		*2*		*0*	

Abbreviations: BrCa= Breast cancer; DCIS=Ductal carcinoma *in situ* ; LCIS = Lobular carcinoma *in situ* ; MR =Medical report (medical record and/or pathology report)

[1] Breast cancer events (any type) are defined by the following ICD10 codes: D05 (includes lobular carcinoma *in situ*, ductal carcinoma *in situ*, other and unspecified *in situ* breast carcinomas), C50 (all invasive breast cancer), and D48.6 (neoplasm of uncertain behavior of breast)

[2] Excludes n=1 where race/ethnicity is missing

[3] Data release 9.0 includes events from follow-up phases that closed **on or before 9/30/2019**, excluding all data from participants who have withdrawn their data from the study (n=3). To be considered a breast cancer event, breast must be the primary site, and the cancer diagnosis must have occurred after the participant's baseline date. Consequently, Data Release 9.0 breast cancer events (any type) exclude a) through c) below:

a) Breast cancers diagnosed prior to participant's baseline date;

b) Breast cancers diagnosed with unknown timing relative to baseline;

c) Unknown primary cancer diagnoses where breast was unlikely to be the primary site, but the possibility could not be ruled out

[4] Includes 'presumed invasive' (i.e. a carcinoma *in situ* with positive nodes)

[5] Ductal only, Lobular only, and Ductal and Lobular exclude inflammatory breast cancer and Paget's

[6] Includes BrCa events with a) tumors with mixed ductal and lobular characteristics and/or b) multiple tumors of different histology (e.g. a ductal only tumor and a lobular only tumor)

[7] Phyllodes is neither ductal nor lobular; it is a sarcoma

[8] Includes: a) Invasive but unknown if ductal or lobular, these are typically self-report without MR confirmation; b) Invasive type "neither ductal nor lobular, not otherwise specified", these are typically self-report without MR confirmation; and c) Per MR, presumed invasive - *in situ* tumor, positive nodes;

[9] *In situ* Ductal only, i*n situ* Lobular only, and i*n situ* Ductal and Lobular exclude *in situ* Paget's

[10] Includes: a) *In situ* but unknown if ductal or lobular; b) *In situ* type "neither ductal nor lobular, not otherwise specified"; these are typically self-report without MR confirmation

[11] Invasiveness unknown includes: a) Unknown if invasive or *in situ* and unknown if ductal or lobular; b) Ductal but unknown if invasive or *in situ*; c) Lobular but unkonwn if invasive or *in situ*; d) Neither ductal nor lobular (other) but not otherwise specified and unknown if invasive or *in situ*; e) Both ductal and lobular but unknown if invasive or *in situ*; f) Per MR, unable to determine if DCIS or atypical ductal hyperplasia

[12] Priority for stage and ER/PR/HER2 status when diagnosis is reported by multiple sources: Medical record>Breast Cancer Follow-Up Form

[13] Medical confirmation of reported breast cancer is generally from medical records but may only include pathology report; National Death Index (NDI) Plus or death certificate confirmation is specific to cases that do not have medical confirmation

[14] For those missing age at diagnosis, age was imputed as the midpoint of the range of plausible ages at diagnosis