Sister Study First DCIS or Invasive Breast Cancer Events [1] and Characteristics: Data Release 8.1 [2]

Characteristic	N	% of cohort
Total Cases	**3,681**	**7.2**

Breast Cancer Subtype for 1st DCIS or Invasive Diagnosis	N	% of cases
Invasive [3]	2,879	78.2
Ductal only (no lobular component) [4]	2,018	54.8
Lobular only (no ductal component) [4]	252	6.8
Ductal and lobular [4,5]	188	5.1
Inflammatory [6]	6	0.2
Paget's [7]	10	0.3
Phyllodes [8]	5	0.1
Other/unknown if ductal or lobular [9]	400	10.9
In situ	792	21.5
Ductal only (no lobular component) [10]	767	20.8
Ductal and lobular [5,10]	17	0.5
Paget's [7]	5	0.1
Other [11]	3	0.1
Invasiveness unknown [12]	10	0.3
Stage at 1st DCIS or Invasive Diagnosis [13]		
0	727	22.3
I	1,674	51.5
II	659	20.3
III	170	5.2
IV	21	0.6
Unknown, missing, cannot stage	*430*	
Estrogen Receptor (ER) Status: 1st DCIS or Invasive Breast Cancer [13]		
Positive	2,671	84.7
Negative	476	15.1
Borderline	5	0.2
Unknown, missing, not performed	*529*	
Progesterone Receptor (PR) Status: 1st DCIS or Invasive Breast Cancer [13]		
Positive	2,207	72.2
Negative	843	27.6
Borderline	7	0.2
Unknown, missing, not performed	*624*	
HER2 Amplification Status: 1st DCIS or Invasive Breast Cancer [13]		
Positive	339	13.0
Negative	2,239	85.6
Borderline	37	1.4
Unknown, missing, not performed	*1,066*	
Combined ER, PR, HER2 Status: 1st DCIS or Invasive Breast Cancer [13]		
Triple negative	265	8.7
Not triple negative	2,795	91.3
Unknown, missing, not performed	*621*	

Characteristic	N (total = 3,681)	% of cases
Diagnosis Confirmation [14]		
Medical	2,997	81.4
NDI Plus or death certificate only	39	1.1
None	645	17.5
Race/Ethnicity [15]		
Non-Hispanic White	3,149	85.6
Non-Hispanic Black	288	7.8
Hispanic	139	3.8
Other	104	2.8
Missing	*1*	
Age (in years) at 1st DCIS or Invasive Breast Cancer [16]		
Under 50	398	10.8
50-64	1,729	47.0
65 and older	1,554	42.2
Menopausal Status at 1st DCIS or Invasive Breast Cancer		
Pre-menopausal	604	16.5
Post-menopausal	3,061	83.5
Unknown	*16*	

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: D05.1 (ductal carcinoma in situ), C50 (all invasive breast cancer), and D48.6 (neoplasm of uncertain behavior of breast)

[2] Data release 8.1 includes events from follow-up phases that closed **on or before 9/23/2018**, 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 8.1 breast cancer events for 1st DCIS or invasive diagnosis exclude a) through c) below:

    a) Breast cancers diagnosed prior to participant's baseline date. [Total (n=57): Invasive ductal (n=31); invasive lobular (n=1), invasive ductal and lobular (n=1), invasive, unknown if ductal or lobular (n=5), and DCIS (n=19);

    b) Breast cancers diagnosed with unknown timing relative to baseline [Total (n=17): Invasive, unknown if ductal or lobular (n=10) and DCIS (n=7)];

    c) Unknown primary cancer diagnoses where breast was unlikely to be the primary site, but the possibility could not be ruled out (n=4)

[3] Includes 'presumed invasive' (i.e. a carcinoma in situ with positive nodes)(n=1)

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

[5] 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)

[6] Inflammatory – Ductal (n=5), Ductal and Lobular (n=1), Lobular(n=0)

[7] Paget's - Invasive: Ductal (n=9), Lobular (n=1); *in situ*: Ductal (n=5)

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

[9] Includes: a) Invasive but unknown if ductal or lobular (n=395), these are typically self-report without MR confirmation; b) Invasive type "neither ductal nor lobular, not otherwise specified" (n=4), these are typically self-report without MR confirmation; and c) Per MR, presumed invasive - in situ tumor, positive nodes (n=1)

[10] *in situ* Ductal only and *in situ* Ductal and Lobular exclude *in situ* Paget's

[11] Excludes all with lobular-only origin

[12] Invasiveness unknown includes: a) Unknown if DCIS or invasive ductal (n=7); and b) Per MR, unable to determine if DCIS or atypical ductal hyperplasia (n=3)

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

[14] 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

[15] Race/ethnicity: precedence given to Hispanic ethnicity first, then Black race, then Other race, then White

[16] For those missing age at diagnosis (n=20), age was imputed as the midpoint of the range of plausible ages at diagnosis