

**Data Dictionary for SWOG S1007 Survival Dataset**

**(NCT01272037-D1-Dataset.csv)**

**PubMed ID: 34914339**

<b>Variable</b>	<b>Definition</b>	<b>Type</b>	<b>Code</b>	<b>Code Definition</b>
STUDYID	SWOG study number (S1007 for all)	Char		
ARMNAME	Treatment arm	Char	Chemo and Endocrine Therapy	Randomized to chemotherapy and endocrine therapy
			Endocrine Therapy Alone	Randomized to endocrine therapy alone
HISPANIC	Reported Hispanic ethnicity	Char	Y	Hispanic
			N	Not Hispanic
			U	Unknown/not reported
SEX	Sex (female for all)	Char	F	Female
ALTPATID	Unique participant identifier	Char		
RACE	Reported race	Char	B	Black or African American
			M	Multiple races
			N	American Indian or Alaska Native
			P	Native Hawaiian or Other Pacific Islander
			S	Asian
			W	White
			U	Unknown or not reported
ELIG	Eligibility indicator	Num	0	Ineligible
			1	Eligible
INCLUDE	Indicator of inclusion in the analysis; participants who were eligible and maintained consent	Num	0	Not included
			1	Included

CHEMO	Indicator of being randomized to the chemotherapy arm	Num	0	No (equivalent to ARMNAME = 'Endocrine Therapy Alone')
			1	Yes (equivalent to ARMNAME = 'Chemo and Endocrine Therapy')
DFSTIME0	Indicator of zero follow-up	Num	0	Any follow-up
			1	No follow-up
PERPROTOCOL	Indicator of acceptance of assigned treatment at time of randomization	Num	0	Did not accept treatment assignment
			1	Accepted treatment assignment
PRESENT_EVENT	First Invasive Disease-Free Survival (IDFS) event type	Num	1	Any distant recurrence
			2	Local/regional recurrence
			3	Opposite breast (recurrence or new primary)
			4	Non-breast new primary
			5	Recurrence not classified
			6	Death not due to recurrence or second primary
SECOND_DISTANT	Indicator of distant recurrence after the first event	Char	N	No
			Y	Yes
NEWREG	Indicator of registration after the study protocol was revised to exclude participants with micrometastases as the only nodal involvement	Num	0	No
			1	Yes

SWOGPF	Zubrod performance status	Num	0	Fully active, able to carry on all pre-disease performance without restriction
			1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work
			2	Ambulatory and capable of self-care but unable to carry out any work activities; up and about more than 50% of waking hours
YEARREG	Year of registration	Num		
RS1425	Oncotype DX Recurrence Score, grouped	Num	0	0-13
			1	14-25
MENOP	Menopausal status	Num	0	Premenopausal
			1	Postmenopausal
SENTONLY	Type of nodal dissection	Num	0	Axillary lymph node dissection with or without sentinel node mapping
			1	Sentinel node biopsy only
CHEMORS	Chemotherapy*Recurrence Score interaction	Num		
DFSIND	Invasive disease-free survival (IDFS) indicator	Num	0	Censored
			1	Recurrence, new primary, or death

DFSTIM	Invasive disease-free survival (IDFS) in days	Num		
SURIND	Overall survival (OS) indicator	Num	0	Censored
			1	Death
SURTIM	Overall survival (OS) in days	Num		
TUMSIZ	Tumor size	Num	1	< 2 cm
			2	>= 2 and < 5 cm
			3	≥ 5 cm
			.	Not Reported
POSNODES	Number of positive lymph nodes	Num	1	1
			2	2
			3	3
PGRPOS	Progesterone receptor (PgR) positivity indicator	Num	0	Negative
			1	Positive
			Blank	Not Reported
LVI	Lymphovascular space invasion indicator	Num	0	No
			1	Yes
EXE	Extranodal extension indicator	Num	0	No/Unknown
			1	Yes
			Blank	Not Reported
HISTOLOGY	Tumor histology	Num	1	Ductal
			2	Lobular
			3	Mixed ductal/lobular
			4	Other
			Blank	Not Reported
OVARSUP	Indicator for ovarian suppression in months 0-6 from randomization and/or months 7-12 from randomization	Num	0	No
			1	Yes
BCT	Breast conserving surgery vs mastectomy indicator	Num	0	Mastectomy
			1	Breast conserving surgery
AGE	Age in years (rounded down)	Num		

AGEDECILE	Age, grouped	Num	1	<40 years
			2	40-49 years
			3	50-59 years
			4	60-69 years
			5	70+ years
DFSTIMYRS	Invasive disease-free survival (IDFS) in years	Num		
RS	Oncotype DX Recurrence Score	Num	Integer from 0-25	
DRFSTIM	Distant relapse-free survival (DRFS) in days	Num		
DRFSIND	Distant relapse-free survival (DRFS) indicator	Num	0	Censored
			1	Distant recurrence or death
DRFSTIMYRS	Distant relapse-free survival (DRFS) in years	Num		
ENDOCAT	Reviewed endocrine therapy type during (Table S2A)	Char	AI	Aromatase inhibitor
			None Reported	None
			OFS and AI	Ovarian function suppression and aromatase inhibitor
			OFS and Tam	Ovarian function suppression and Tamoxifen
			OFS/Tam/AI	Ovarian function suppression and Tamoxifen and aromatase inhibitor
			Other	Other
			Tam	Tamoxifen
			Tam and AI	Tamoxifen and aromatase inhibitor
			.	Not Reported

CHEMOTYPE	Reviewed chemotherapy type (Table S2A)	Char	Anthra and Taxane	Anthracycline and taxane
			Anthra w/o Taxane	Anthracycline without taxane
			None	None
			Other	Other
			Taxane & Cyclophosphamide	Taxane and cyclophosphamide
			.	Not Reported
GRADE2GRP	Tumor histologic grade, grouped	Num	1	Grade 1 (Low)
			2	Grade 2-3 (Intermediate or High)
TUMSIZ2GRP	Tumor size, grouped	Num	1	<2 cm
			2	≥ 2 cm
POSNODES2GRP	Number of positive lymph nodes, grouped	Num	1	1 positive node
			2	2-3 positive nodes
CLINTOX	Evaluable for toxicity	Char	Y	Evaluable
			N	Not evaluable
			P	Pending
GRADE	Tumor histologic grade	Char	G1	Grade 1 (Low)
			G2	Grade 2 (Intermediate)
			G3	Grade 3 (High)
			GX	Unknown
			Blank	Not Reported
MENSTSTOP	Indicator for no periods in months 0-6 and months 7-12	Num	1	No period within 12 months of randomization
			0	Experienced periods during 12 months of randomization
			.	Not Reported

#### NOTES

- 1) Unless otherwise specified, ‘.’ = missing data.

- 2) Analyses only include participants who maintained consent to stay on the study and were eligible (INCLUDE=1), while reporting of baseline characteristics includes all eligible participants (ELIG=1).
- 3) Data can be used to approximate published study findings, but exact reproduction of previous manuscripts may not be possible in some cases (e.g., when data must be modified for de-identification purposes or have undergone further data cleaning). In this submission, baseline characteristics data are more up-to-date than the publication (e.g. Tables S1 and S6). In addition, some denominator participant counts in the survival results may differ due to the ways different programs handle participants with zero follow-up time. This does not affect the survival estimates or hazard ratios.
- 4) Table 2 in the publication has a repeated row, resulting in some values being misaligned. This is being corrected directly with the journal.
- 5) The patient with a negative invasive disease-free survival time was found to have bone metastasis prior to randomization, and subsequently did not receive therapy as planned. Therefore, the DFSTIM less than 0 indicates disease recurrence prior to registration. This patient was not eligible and was not included in analysis (ELIG=0, INCLUDE=0).