

## E2108 Primary Analysis Clinical Data Description

The clinicaltrials.gov protocol identifier: NCT01242800

This data dictionary is for the NCT01242800-D1 dataset.

The data in the spreadsheet are the clinical analysis data for the primary manuscript on E2108 results, published in

Khan SA, Zhao F, Goldstein LJ, Cella D, Basik M, Golshan M, Julian TB, Pockaj BA, Lee CA, Razaq W, Sparano JA, Babiera GV, Dy IA, Jain S, Silverman P, Fisher CS, Tevaarwerk AJ, Wagner LI, Sledge GW. Early Local Therapy for the Primary Site in De Novo Stage IV Breast Cancer: Results of a Randomized Clinical Trial (E2108). *J Clin Oncol*. 2022 Mar 20;40(9):978-987.PMCID: PMC8937009.

The data reflect the study database as of December 10, 2019. It includes all patients registered to the Step 1 of the trial.

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).

### Data File Description

Field	Variable Name	Description	Coding	Note
1	uid	Patient identifier		De-identified patient ID
2	step	Indicator for whether the patient was randomized to step 2	0=Off study at step 1, 1=Randomized to step 2	
3	elig_step1	Met eligibility criteria for step 1	0=No, 1=Yes	
4	age1	Age at study registration (years) to step 1	Continuous variable	
5	gender	Sex	1=Male, 2=Female	
6	race	Race	1=White, 2=Black, 3=Other	
7	ethn	Ethnicity	1=Non-Hispanic, 2=Hispanic	
8	menopausal_stat1	Menopausal status at step 1	1=Postmenopausal, 2=Premenopausal, 3=Perimenopausal	

Field	Variable Name	Description	Coding	Note
9	subtype	Tumor subtype	1=HR-positive and HER2-negative, 2=Triple-negative, 3=HER2-positive	
10	metasite1	Metastatic sites at step 1	1=Visceral only, 2=Bone only, 3=Soft tissue only, 4=Multiple sites	
11	oligo	Oligomets	0=No, 1=Yes	The 384 patients who had assessable site of disease had value for this variable.
12	stage	Primary tumor stage	1=T1-3 and N0-1, 2=T4 and/or N2-3	
13	invasion_yn	Direct invasion of tumor into skin	1=No, 2=Yes	
14	nod_prsnt_yn	(Breast) Skin nodules present	1=No, 2=Yes	
15	tum_loc	Tumor mobility	1=Free, 2=Attached to fascia, 3=Chest wall fixation	
16	OSTdata	Indicator whether patient reported optimal systemic therapy (OST) data	0=No, 1=Yes	
17	OST_timing	Began OST prior to study registration	0=No, 1=Yes	Only applicable for the 374 patients who reported OST data.
18	OST_group	OST treatment regimen combination	1=Chemo alone, 2=Endocrine therapy alone, 3=Both endocrine and chemotherapy, 4=Chemotherapy with HER2-directed therapy	Only applicable for the 374 patients who reported OST data.
19	OST_week	OST treatment duration in weeks after study registration	Continuous variable	Only applicable for the 374 patients who reported OST data.
20	off_tx_reas	Reason for discontinuing protocol therapy at Step 1	1=Disease progression, 2=Patient withdrawal, 3=Unsuitable for surgery, 4=Death on study, 5=Ineligible for random assignment, 99=Other reason	Only the 134 patients who went off study after step 1 had value for this variable.
21	surv_s1	Overall survival (OS) status for patients off treatment at Step 1	0=Censored, 1=Died	As of database on 12/10/2019, only the 134 patients who went

Field	Variable Name	Description	Coding	Note
				off study after step 1 had value for this variable.
22	surv_m1	OS time in months for patients off treatment at Step 1	Continuous variable	As of database on 12/10/2019, only the 134 patients who went off study after step 1 had value for this variable.
23	cod_step1	Cause of death	2=Breast cancer, 3=Other disease, 9=Cause unknown	As of database on 12/10/2019, only the patients who went off study at step 1 and died had value for this variable.
24	trtm	Treatment arm at Step 2	A=Continued systemic therapy, B=Early local therapy	The 256 patients who were randomized to step 2 had value for this variable

Notes:

- Unless otherwise specified, blank values indicate missing data.
- In the analysis, months were calculated as number of days/30.4375.
- The data in this dataset that are reported in Supplemental Table 1 are further cleaned since publication.