EA1131 Primary Analysis Clinical Data Description

This data dictionary is for NCT02445391-D1 dataset.

The data in the spreadsheet are the clinical analysis data for the primary manuscript on EA1131 results, published in

> Mayer IA; Zhao F; Arteaga CL; Symmans WF; Park, Ben H; Burnette, Brian L; Tevaarwerk, Amye J; Garcia, Sofia F; Smith, Karen L; Makower, Della F; Block, Margaret; Morley, Kimberly A; Jani, Chirag R; Mescher, Craig; Dewani, Shabana J; Tawfik, Bernard; Flaum, Lisa E; Mayer, Erica L; Sikov, William M; Rodler, Eve T; Wagner, Lynne I; DeMichele, Angela M; Sparano JA; Wolff AC; Miller KD. Randomized Phase III Postoperative Trial of Platinum-Based Chemotherapy Versus Capecitabine in Patients With Residual Triple-Negative Breast Cancer Following Neoadjuvant Chemotherapy: ECOG-ACRIN EA1131. *J Clin Oncol*. 2021 Aug 10;39(23):2539-2551.

The data reflect the study database as of April 7, 2021. All patients who were randomized to one of the three arms were included in the database. Accrual was open to arm A (observation) and arm B (platinum) prior to Amendment #3, and accrual to arm A was permanently closed and accrual to arm C (capecitabin) was opened in Amendment #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).

Data File Description

Field	Variable Name	Description	Coding	Note
1	uid	Patient identifier		De-identified patient ID, not ECOG case ID
2	studyperiod	Enrollment period	1=enrolled prior to Amendment #3 2=enrolled after Amendment #3 3=enrolled after Amendment #7	Accrual open to Arm A and Arm B prior to activation of Amendment #3. Accrual to arm A was permanently closed and accrual to arm C was opened in Amendment #3. Intrinsic subtype was removed

Field	Variable Name	Description	Coding	Note
				as a stratification factor in Amendment #7.
3	indicator	Whether the case was included in the primary analysis	0=no, 1=yes	Basal subtype patients who were randomized to arm B and arm C after Amendment #3 were the primary analysis population
4	subtype	Intrinsic subtype	0=non-basal subtype, 1=basal subtype	Intrinsic subtype was available for 401 patients as of the data cutoff date, it was pending for 14 patients
5	treated	Received protocol therapy	0=no, 1=yes	
6	trtdata	Whether treatment data were submitted	0=no, 1=yes	391 patients who were randomized to arms B and C and started protocol therapy had value for this variable. Treatment data were not submitted yet for 7 patients as of the data cutoff date
7	ontrt	Indicator whether patients were receiving protocol treatment at the data cutoff date	1= yes, on protocol treatment at data cutoff date	
8	trtm	Treatment arm	A= Arm A, B= Arm B, C=Arm C	Arm A=observation, Arm B= platinum, Arm C=capecitabin
9	agnt_choice	Platinum agent choice	String variable	
10	elig	Eligibility	0=no, 1=yes	
11	var1	Stratification factor -Clinical staging	1= II, 2=III	
12	var2	Stratification factor – Residual cancer burden after neoadjuvant chemotherapy	1=1-3cm, 2=>3cm	
13	var3	Stratification factor - Planned platinum agent	1=cisplatin, 2=carboplatin	

Field	Variable Name	Description	Coding	Note
14	var4	Stratification factor- anthracycline exposure in the neoadjuvant setting	1=no, 2=yes	
15	var5	Stratification factor - administration of radiotherapy	1=no, 2=yes	
16	var6	Stratification factor – intrinsic subtype	1=non-basal, 2=basal	Intrinsic subtype was added as a stratification factor in Amendment #3 and was removed later in Amendment #7. Before amendment #3 (activated in June 2016), only basal subtype TNBC patients were eligible for the study. After amendment #7 (activated in November 2020), all patients were eligible as long as they submitted their specimen for PAM50 analysis regardless of the assay results (eg, assay failure, basal subtype, non-basal subtype)
17	surv_s	Overall survival (OS) status	0=censor, 1=died	As of database on 4/7/2021
18	surv_m	OS time in months	Continuous variable	As of database on 4/7/2021
19	dfs_s	Disease-free survival (DFS) status	0=censor, 1=DFS event	As of database on 4/7/2021
20	dfs_m	DFS time in months	Continuous variable	As of database on 4/7/2021
21	rfs_s	Recurrence-free survival (RFS) status	0=censor, 1=RFS event	As of database on 4/7/2021
22	rfs_m	RFS time in months	Continuous variable	As of database on 4/7/2021
23	recur	Disease recurrence	0=no, 1=yes	As of database on 4/7/2021
24	recur_dist	Distant recurrence	0=no, 1=yes	As of database on 4/7/2021
25	recur_lr	Locoregional recurrence	0=no, 1=yes	As of database on 4/7/2021
26	age	Age at study entry (years)	Continuous variable	
27	sex	Sex	String variable	
28	race	Race	1=White, 2=Black, 3=Asian, 4=unknown	

Field	Variable Name	Description	Coding	Note
29	ethn	Ethnicity	1=Non-Hispanic, 2=Hispanic, 3=unknown	
30	ecogps	ECOG performance status	0=0, 1=1	
31	prior_rt_yn	Prior radiation therapy	String variable	
32	prior_tax_chemo	Prior neoadjuvant taxane	String variable	
33	prior_anth_chemo	Prior neoadjuvant anthracycline	String variable	
34	prior_adj_other	Other prior neoadjuvant therapy	String variable	
35	extnsv_surg	Most extensive surgery	String variable	
36	surg_rand_gap	Days between surgery and randomization	Integer	
37	rand_trt_gap	Days between randomization and initiation of protocol treatment	Integer	
38	histology	Histology of primary cancer	1=invasive ductal carcinoma, 2=other	
39	histol_grade	Histologic grade	String variable	
40	prim_tum_size	Primary tumor size	Continuous variable	
41	tstage	T stage	1= cT1, 2= cT2,3= cT3, 4= cT4	
42	nstage	N stage	0=cN0, 1=cN1, 2=cN2, 3=cN3, 9=cNX	
43	resid_ca_brdn	Residual cancer burden in cm	Continuous variable	
44	ypstage	Pathological yp stage	1=stage I, 2=stage II, 3= stage III	
45	ptstage	Pathological T stage	String variable	
46	ln_involv_yn	Lymph node involvement	String variable	
47	sentinel_bx_yn	Sentinel node biopsy	String variable	
48	axil_diss_yn	Axillary node dissection	String variable	
49	totalcycle	Total number of cycles of protocol treatment	Integer	384 patients who started protocol therapy had the value for the variable, of them 368 patients had discontinued protocol therapy as of the data cutoff date

Field	Variable Name	Description	Coding	Note
50	offtrt	Indictor for whether patient went off treatment as of the data cutoff date	1=yes	368 patients started protocol therapy and had discontinued therapy as of the data cutoff date. Other patients either never started protocol therapy or were still on treatment as of the data cutoff date
	off_tx_reas	Reason for discontinuing protocol therapy	String variable	Only the 368 patients who started and discontinued protocol therapy had value for this variable.
	second	Second primary cancer	0=no, 1=yes	As of database on 4/7/2021
	cod	Cause of death	2=disease related 3=other disease 99=not known	98 patients who died had value for this variable
	daygap	Days between surgery and initiation of protocol treatment	Integer	
	dose_mod	Any dose modification during treatment	String variable	383 patients who started protocol therapy on arms B and C and reported dose modification data had value for this variable
	dosereduce	Indicator for any dose reduction during treatment	1=yes	77 patients on arm B reported dose reduction and had value for this variable

Notes:
- Unless otherwise specified, blank values indicate missing data.
- In the analysis, months were calculated as days/30.4375