

**Data Dictionary for SWOG S1314 Dataset  
(NCT02177695-D1-Dataset.csv)**

**PMID: 33568346**

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).

The dataset for this phase 2 trial in the Archive does not contain the toxicity data presented in Table 4.

Patient data set (N=227 rows)	Description	Coded values
ALTPATID	Unique patient identifier	(alphanumeric)
ARMNAME	Treatment arm to which the patient was randomized	GC+CYST = gemcitabine + cisplatin + cystectomy DDMVAC+CYST = dose-dense methotrexate-vinblastine-adriamycin-cisplatin + cystectomy
STRAT1	Baseline clinical T stage; stratification factor;	Clinical Stage: T2 Clinical Stage: T3 or T4a
STRAT2	Baseline Zubrod performance status; stratification factor;	Performance status: 0 Performance status: 1
AGE	Age (in years) at the time of registration	(numeric)
SEX	Sex, as reported at the time of randomization	M = Male F = Female
RACE	Race, as reported at the time of randomization	B = Black or African American N = American Indian or Alaska Native S = Asian W = White U = Unknown
PRIM_ANALYSIS	Indicator for which patients were included in the primary COXEN analysis (N=167)	1= included in primary COXEN analysis 0= excluded from primary COXEN analysis
ITT_ANALYSIS	Indicator for which patients were included in the secondary ITT analysis (N=227)	1= included in the ITT analysis 0= excluded from ITT analysis
ADEQUATE_TISS	Result of tissue evaluation for adequacy for COXEN analysis.	Y = tissue was adequate N = tissue was inadequate missing = ineligible patients did not have their tissue evaluated for adequacy
ANY_TRT	Receipt of any amount of protocol treatment	Y = received any amount of protocol treatment N = Did not receive any protocol treatment
GFS	Receipt of growth factor support during chemotherapy (Filgrastim or Pegfilgrastim)	1 = growth factor support received 0 = no growth factor support received
ADEQUATE_CHEMO	Receipt of at least 3 cycles of protocol-specified cisplatin-based chemotherapy; required to be considered evaluable for the primary COXEN analysis	Y = patient received at least three cycles of chemotherapy (or progressed while receiving chemotherapy) N = patient received less than 3 cycles of chemotherapy or no chemotherapy
CYSTECTOMY	Receipt of cystectomy within 100 days following last dose of chemotherapy, or progression prior to expected cystectomy; required to be considered evaluable for the primary COXEN analysis	Y = patient underwent a cystectomy within 100 days after chemotherapy or progressed prior to the expected cystectomy(the latter were managed as nonresponders) N = patient did not undergo cystectomy, or underwent cystectomy outside of 100 days after chemotherapy
PATHSTGT	Pathologic T Stage of cystectomy tissue as determined by the local pathologist; missing for pts who did not get cystectomy	PTX PT0 PTIS PTA PT1 PT2 PT2A PT2B PT3 PT3A PT3B PT4A

PATHSTGN	Pathologic N Stage of cystectomy tissue as determined by the local pathologist; missing for pts who did not get cystectomy	PNX PN0 PN1 PN2 PN3
PATHSTGM	Pathologic M Stage of cystectomy tissue as determined by the local pathologist; missing for pts who did not get cystectomy	M0 M1
NLYMPHN	Number of lymph nodes examined	(numeric) "." = missing data
NPOSNOD	Number of positive lymph nodes	(numeric) "." = missing data
PATH_RESP	Pathologic response determined from cystectomy specimen; if no cystectomy, coded as a non-responder (NR);	CR = Complete Response (pT0) PR = Partial Response or Downstaged (>pT0, but <= pT1) NR = No Response
OS_IND	Overall survival indicator	1=event 0=censored
OS_TIM	Days from study registration to death, or date of last contact	(numeric)
PFS_IND	Progression-free survival indicator	1=event 0=censored
PFS_TIM	Days from study registration to progression/recurrence or death, censored at date of last contact	(numeric)
TT_OFFTX	Days from study registration to off chemo	(numeric)
TT_CYST	Days from study registration to cystectomy; missing for pts who did not get cystectomy	(numeric)
GC_FAV	Favorable GC GEM score;	1 = favorable 0 = not favorable missing = not evaluable for the primary COXEN analysis and therefore did not have COXEN scores generated
DDMVAC_FAV	Favorable ddMVAC GEM score;	1 = favorable 0 = not favorable missing = not evaluable for the primary COXEN analysis and therefore did not have COXEN scores generated
GC_FAV_DUP	Favorable GC GEM score from batch 2 (duplicate); Seven patients had their samples run in two batches with GEM scores calculated and dichotomized for both GC and ddMVAC (14 determinations).	1 = favorable 0 = not favorable missing = not included in the duplicate batch
DDMVAC_FAV_DUP	Favorable ddMVAC GEM score from batch 2 (duplicate); Seven patients had their samples run in two batches with GEM scores calculated and dichotomized for both GC and ddMVAC (14 determinations).	1 = favorable 0 = not favorable missing = not included in the duplicate batch
GEO	Patient's molecular data is uploaded in GEO The record ID in GEO for study S1314: GSE244266	1=yes 0=no
TCGAGRP	TCGA Classifier Label – RNA-based molecular subtypes from TCGA	(character) "" = missing data
TCGA	TCGA Classifier Number	(numeric) "" = missing data
CONSENSUS	Consensus Classifier Label – RNA-based molecular subtypes from Consensus group	(character) "" = missing data
MDA	MDA Classifier Number	(numeric) "" = missing data
MDAGRP	MDA Classifier Label – RNA-based molecular subtypes from MD Anderson	(character) "" = missing data
POOR_QUAL	After the primary analysis was conducted, the Affymetrix expression data was deemed to be poor quality as assessed by the Director of Genomics for Greenberg Bladder Cancer Institute	1=poor quality 0=adequate quality