

NRG Oncology/RTOG 9804 Data Dictionary
NCTN Data Archive
(n=636)

#	Variable	Description	Coding for categorical/Min-max for continuous
1	study_no	Study number	RTOG-9804
2	pt_id	De-identified patient ID	Continuous Min-max = 1-636
3	rx	Assigned treatment	1 = Observation ± tamoxifen 2 = Radiotherapy ± tamoxifen
4	case_status	Case Status	1 = Eligible 2 = No protocol treatment 3 = Ineligible 4 = Withdrew consent
5	analysis_flag	Analyzed for the primary endpoint?	1 = No (n=51) [case_status=3,4] 2 = Yes (n=585) [case_status=1,2]
6	reason_for_exclusion	Reason for Exclusion (for ineligible/withdrew consent patients)	Character variable of length 100 Reason is not blank (n=51) [analysis_flag=1] Reason is blank (n=585) [analysis_flag=2]
7	analysis_flag_sensitivity	Randomly allocated with follow-up information?	1 = No (n=7) 2 = Yes (n=629) Note: Patients with analysis_flag_sensitivity = yes were used for a sensitivity analysis of the primary endpoint.
8	treated	Started protocol treatment out of patients analyzed for the primary endpoint?	1 = No (n=6) 2 = Yes (n=579) . = Not applicable (n=51) [analysis_flag=1]
9	age	Age (years)	Continuous Min-max = 34-85 . = Not applicable (n=7; analysis_flag_sensitivity=1)
10	agefm	Age (years)-Categorized	1 = < 50 2 = ≥ 50 . = Not applicable (n=7; analysis_flag_sensitivity=1)
11	race	Race	1 = White 2 = African American 3 = Hispanic 4 = Native Hawaiian or other Pacific Islander 5 = Asian 6 = American Indian or Alaska Native 7 = Other . = Not applicable (n=7; analysis_flag_sensitivity=1)
12	tamoxifen_use	Intention to use tamoxifen?	1 = No 2 = Yes . = Not applicable (n=7; analysis_flag_sensitivity=1)
13	tum_location	Tumor location	1 = Left breast 2 = Right breast . = Not applicable (n=7; analysis_flag_sensitivity=1)
14	nuclei_grade	Nuclei grade	1 = NG1 2 = NG2 . = Not applicable (n=7; analysis_flag_sensitivity=1)
15	mammsize	Mammographic size of primary tumor, cm	1 = ≤ 1cm 2 = > 1cm . = Not applicable (n=7; analysis_flag_sensitivity=1)
16	margins	Final microscopic margins, mm	1 = ≥ 3mm - 9mm 2 = ≥ 10mm

#	Variable	Description	Coding for categorical/Min-max for continuous
			3 = Negative by negative re-excision 999 = Missing (n=1) . = Not applicable (n=7; analysis_flag_sensitivity=1)
17	pathsize	Largest dimension of DCIS lesion on pathology slide, cm	Continuous Min-max = 0.0-2.5 999=Missing (n=2) . = Not applicable (n=7; analysis_flag_sensitivity=1)
18	local_failure	Local failure status	0 = censored 1 = event (defined as biopsy-proven invasive or DCIS recurrence) 2 = competing event (death) . = Not applicable (n=7; analysis_flag_sensitivity=1)
19	local_failure_years	Local failure time (years)	Time from the date of randomization to the date of local failure or last follow-up/death.
20	local_failure_histology	Histology of local failure	0 = No local failure 1 = Invasive 2 = Non-invasive . = Not applicable (n=7; analysis_flag_sensitivity=1)
21	survival	Overall survival status	0 = censored 1 = event (death) . = Not applicable (n=7; analysis_flag_sensitivity=1)
22	survival_years	Overall survival time (years)	Time from the date of randomization to the date of death or last follow-up.
23	distant_metastasis	Distant metastasis status	0 = censored 1 = event (progression of disease beyond the treated breast and regional nodes) 2 = competing event (death) . = Not applicable (n=7; analysis_flag_sensitivity=1)
24	distant_metastasis_years	Distant metastasis time (years)	Time from the date of randomization to the date of distant metastasis or last follow-up/death.
25	contralateral_failure	Contralateral breast failure status	0 = censored 1 = event (contralateral breast failure) 2 = competing event (death) . = Not applicable (n=7; analysis_flag_sensitivity=1)
26	contralateral_failure_years	Contralateral breast failure time (years)	Time from the date of randomization to the date of contralateral breast failure or last follow-up/death.
27	mastectomy_failure	Salvage mastectomy failure status	0 = censored 1 = event (defined as removal of the study breast) 2 = competing event (death) . = Not applicable (n=7; analysis_flag_sensitivity=1)
28	mastectomy_failure_years	Salvage mastectomy failure time (years)	Time from the date of randomization to the date of mastectomy or last follow-up/death.
29	mastectomy_type	Type of salvage mastectomy	0 = No salvage mastectomy 1 = Ipsilateral 3 = Bilateral 9 = Unknown . = Not applicable (n=7; analysis_flag_sensitivity=1)
30	disease_free_survival	Disease-free survival status	0 = censored 1 = event (death, local_failure, contralateral_failure, distant_metastasis, mastectomy_failure) . = Not applicable (n=7; analysis_flag_sensitivity=1)
31	disease_free_survival_years	Disease-free survival time (years)	Time from the date of randomization to the date of failure/death or last follow-up.
32	tamoxifen_given	Tamoxifen received?	1 = No

#	Variable	Description	Coding for categorical/Min-max for continuous
			2 = Yes .= Not applicable (n=7; analysis_flag_sensitivity=1)
33	acute_max_overall_grade	Worst overall grade per patient for acute treatment toxicities (≤ 90 days from treatment start) [<i>Graded using National Cancer Institute Common Terminology Criteria, version 2.0.</i>]	0 = grade 0 1 = grade 1 2 = grade 2 3 = grade 3 4 = grade 4 .= Not applicable (n=7; analysis_flag_sensitivity=1)
34	has_late_toxicity	Evaluable for late RT toxicities?	1 = No 2 = Yes .= Not applicable (n=7; analysis_flag_sensitivity=1)
35	late_max_overall_grade	Worst overall grade per patient for late radiotherapy toxicities (> 90 days from treatment start) [<i>Graded using the RTOG/European Organization for Research and Treatment of Cancer Late Radiation Morbidity Scheme.</i>]	0 = grade 0 1 = grade 1 2 = grade 2 3 = grade 3 999 = Did not receive protocol treatment or did not have late RT toxicity information (n=322) .= Not applicable (n=7; analysis_flag_sensitivity=1)