

Variables derived or obtained from medical record abstract (ERS) data

Dataset: MA_VARS.sas7bdat

All variables are defined for the first primary invasive breast tumor unless otherwise specified.

One record per subject

Variable	Description	Comment
STUDYID	CBCS study ID	
DX_DATE	Date of diagnosis of first primary invasive breast cancer	Obtained from P3MA file (ERS). Needs Security Level 3 in IRB to request this variable.
DT_LST_CONT	Date of last contact	For deceased subjects, this is the date of death.
LATERALITY	Laterality R = Right L = Left B = Bilateral	Obtained from P3MA (ERS) file. Bilateral cases are defined as: 1) Participants who are diagnosed with simultaneous bilateral breast cancer (two separate primary invasive tumors discovered at the same time). 2) Participants who are diagnosed with a first invasive breast cancer and a second primary breast cancer (invasive or in situ) within six months of the date of diagnosis of the first invasive tumor.

BILATERAL_GP	<p>Bilateral cases group</p> <p>1 = Invasive in one breast and CIS in the other breast</p> <p>2 = Invasive tumors in both breast</p>	<p>Only available for 67 bilateral cases (LAT=3), missing data for others.</p> <p>Note: Six cases with ipsilateral invasive tumors have missing data.</p>
PRIMARY_HX	<p>History of other primary cancer diagnosed simultaneously or after the diagnosis of primary breast cancer</p> <p>0 = No other primary cancer</p> <p>1 = Bilateral, CIS tumor diagnosed <=6 months of the 1st invasive breast tumor</p> <p>2 = Ipsilateral invasive breast tumor, diagnosed <= 6 months of the first invasive breast tumor</p> <p>3 = Bilateral invasive breast tumor, diagnosed <=6 months of the first invasive breast tumor</p> <p>4 = Multiple invasive tumors, ipsilateral & bilateral diagnosed <= 6 months of the 1st invasive breast tumor</p> <p>5 = Second invasive breast primary diagnosed >6 months after the first invasive primary breast tumor</p> <p>6 = CIS tumor diagnosed >6 months after the first invasive breast tumor</p> <p>7 = Other (not breast) primary tumor</p> <p>8 = Invasive breast primary diagnosed >6 months after the first primary + other primary tumor</p> <p>9 = Ipsilateral invasive breast tumor (<= 6 months) + other primary tumor</p> <p>10 = CIS (<=6 months) + Other primary tumor</p> <p>11 = bilateral invasive (<=6 months) + other</p>	<p>This variable will get updated as we follow the cases over time.</p> <p>Note: as of the date this document, most of the second invasive breast primary (category 5) tumor is in the contralateral breast. For 2 cases, the 2nd breast primary is in the same breast.</p>

SIZE	Tumor size (mm) 998 = inflammatory 999 = unknown	Obtained from P3MA (ERS) file. TBD for cases with multiple breast tumors.
ESTSIZE	Tumor size 1 = ≤ 2 cm 2 = >2 – 5 cm 3 = >5 cm	Recoded from SIZE. TBD for cases with multiple breast tumors.
AJCC_GRP	AJCC Stage 1 = Stage I 1A = Stage IA 1B = Stage IB 2A = Stage IIA 2B = Stage IIB 3A = Stage IIIA 3B = Stage IIIB 4 = Stage IV 88 = Not applicable 99 = Unknown	Obtained from P3MA (ERS) file. 1 = Stage I (diagnosed before 1/1/2010) 1A = Stage IA (diagnosed 2010 and beyond) 1B = Stage IB (diagnosed 2010 and beyond) TBD for cases with multiple breast tumors.
STAGE	AJCC Stage 1 = Stage I 2 = Stage II 3 = Stage III 4 = Stage IV	Recoded from AJCC_GRP 1 = 1, 1A, 1B 2 = 2A, 2B 3 = 3A, 3B 4 = 4 TBD for cases with multiple breast tumors.

ND_POS	<p>Number of positive nodes</p> <p>00 = All nodes examined negative</p> <p>01-89 = 1-89 nodes are positive</p> <p>90 = 90 or more nodes are positive</p> <p>95 = Positive aspiration of lymph nodes performed</p> <p>97 = Positive nodes documented, number not specified</p> <p>98 = No nodes examined</p> <p>99 = Unknown</p>	<p>Obtained from P3MA (ERS) file.</p> <p>TBD for cases with multiple breast tumors.</p>
NODES_MALIG	<p>Number of nodes positive for malignancy</p> <p>00 = All nodes examined negative</p> <p>01-89 = 1-89 nodes are positive</p> <p>90 = 90 or more nodes are positive</p> <p>95 = Positive aspiration of lymph nodes performed</p> <p>97 = Positive nodes documented, number not specified</p> <p>98 = No nodes examined</p> <p>99 = Unknown</p>	<p>Recoded from ND_POS and N_STAGE.</p> <p>This variable is almost identical with ND_POS except the value 98 for some cases.</p> <p>If N_STAGE=0 (no lymph node mets) and ND_POS=98 (nodes not examined), assume all nodes negative (NODES_MALIG=0). This is consistent with the definition in CBCS 1 & 2.</p>
NODESTAT	<p>Node status</p> <p>1 = Positive</p> <p>2 = Negative</p>	<p>Recoded from ND_POS and N_STAGE.</p> <p>Positive is defined as one of the followings:</p> <ol style="list-style-type: none"> 1) Number of nodes positive for malignancy >0 2) Staging - Lymph node metastasis <p>If a case has multiple tumors, count as positive if any tumor is node positive.</p> <p>Note: For some clinically staged cases, NODESTAT is positive but ND_POS=0. These cases were considered node positive at time of diagnosis based on imaging studies, or palpation by managing physician. At the time of diagnosis, no nodes were taken. However after the chemotherapy subsequent surgery in which nodes are taken that may not show any positive ones, which is when 0 is entered under ND_POS.</p>

GRADE	<p>Combined grade</p> <p>1 = Well differentiated</p> <p>2 = Moderately differentiated</p> <p>3 = Poorly differentiated</p> <p>9 = not determined</p>	<p>Obtained from P3MA file.</p> <p>This is different from the CGRADE (combined grade) variable from the Centralized Histopathology Review. The CGRADE variable is the preferred one to use in analysis.</p> <p>TBD for cases with multiple breast tumors.</p>
SYS_SURG_SEQ	<p>Systemic surgery sequence</p> <p>0 = No systemic therapy and/or surgical procedures</p> <p>2 = Systemic therapy before surgery</p> <p>3 = Systemic therapy after surgery</p> <p>4 = Systemic therapy both before and after surgery</p> <p>5 = Intraoperative systemic therapy</p> <p>6 = Intraoperative systemic therapy with other therapy adm before or after</p> <p>9 = Sequence unknown</p>	<p>Variable is not finalized because medical abstract of treatment data is ongoing.</p>

FIRST_RX	<p>Type of the first treatment (first course)</p> <p>0 = None 1 = Chemotherapy 2 = Hormone 3 = Radiation 4 = Other 5 = Surgery 6 = Chemotherapy + Hormone 7 = Chemotherapy + Other 8 = Hormone + Other 9 = Chemotherapy + Radiation 10= Biologics 11= Chemo + Biologics 12= Chemo + Biologics + Other 13= Chemo + Biologics + Hormone 14= Chemo + Radiation + Biologics + Other 15= Hormone + Surgery</p>	<p>A subject may have received several first course treatments over time. This is the earliest type of treatment the subject has received.</p> <p>For example:</p> <p>1. If subject had surgery on 6/16/2008 and chemotherapy on 7/22/2008, then FIRST_RX is coded as 5 (surgery).</p> <p>2. If subject started both chemotherapy and hormone therapy on 3/13/2009; had surgery on 7/17/2009, then FIRST_RX is coded as 6 (Chemo + Hormone).</p>
FIRST_RX_DATE	First treatment date (first course)	The earliest treatment date from these treatments: surgery, chemotherapy, radiation, hormone, biologics, and other treatment.
RX_DAYS	Days to first treatment (first course)	<p>Note: definition different from Sasha McGee's paper.</p> <p>Calculated from first treatment date and diagnosis date: (FIRST_RX_DATE – DX_DATE).</p>
RX_DELAY	<p>First course treatment delay</p> <p>1 = 30 days or less 2 = more than 30 days 3 = no treatment</p>	<p>Derived from RX_DAYS. Three cases with no first course treatment assigned in group 3.</p> <p>One subject with no first course treatment but had subsequent treatment is classified as “no treatment” for this variable.</p>

FIRST_SURG_DATE	First surgery treatment date (first course)	<p>Surgery is defined as: SURG (code 20-90) and/or SLNS (code 3-7)</p> <p>For bilateral cases, if surgery dates are different, take the earliest date.</p>
SURG	Surgery code for the first surgery	<p>Some bilateral cases may have different surgery performed on both breasts on the same date. This variable only captures surgery on one of the breasts. Use the surgery dataset to get more detailed surgery info.</p> <p>See list of codes in Codebook. This could have missing values.</p>
SLNS	<p>Scope of lymph node surgery (for the first surgery)</p> <p>0 = None</p> <p>1 = Biopsy or aspiration of regional lymph node, NOS</p> <p>2 = Sentinel lymph node biopsy</p> <p>3 = Number of regional nodes removed unknown or not stated, regional lymph nodes removed, NOS</p> <p>4 = 1-3 lymph nodes removed</p> <p>5 = 4 or more regional lymph nodes removed</p> <p>6 = Sentinel node biopsy and code 3, 4, or 5 at the same time, or timing not stated</p> <p>7 = Sentinel node biopsy and code 3, 4, or 5 at different times</p> <p>9 = Unknown or not applicable</p>	<p>Some bilateral cases may have different surgery performed on both breasts on the same date. This variable only captures surgery on one of the breasts. Use the surgery dataset to get more detailed surgery info.</p>

SURGERY_TYPE	<p>Type of surgery (first course)</p> <p>1 = Mastectomy 2 = Breast conserving surgery 3 = No surgery</p>	<p>Derived from all first course surgeries. If a case has multiple breast tumors, derived variable from all tumors.</p> <p>Mastectomy: if subject has any first course SURG=30-80.</p> <p>BCS: if subject has any first course SURG=20-24 and/or lymph node surgeries SLNS (code 3-7) and has no SURG=30-80.</p> <p>No surgery: if subject has no surgery data or SURG=0-19 and has no SURG>19.</p>
RSN_NO_SURG	<p>Reason no surgery (first course)</p> <p>0 = Cancer directed surgery performed 1 = Cancer directed surgery not performed, not planned part of rx 2 = CD surgery not recommended, contraindicated due to other condition 5 = Cancer directed surgery not performed, patient expired before surgery 6 = Rx recommended, not performed, no reason indicated in record 7 = Patient or guardian refused surgery 8 = Surgery recommended, unknown if done 9 = Unknown</p>	

FIRST_RAD_DATE	First radiation start date (first course)	Missing if no first course radiation treatment. For bilateral cases, if radiation was administered to one side of the breast, take the date of radiation. If radiation was administered to both breasts, take the earliest radiation date. Radiation data from the tumor with the earliest radiation date will be used in this file.
RAD_END_DT	Radiation end date	End date associated with FIRST_RAD_DATE.
REG_DOSE_CGY	Radiation dose cGy (4-digit field) 8888 = radiation administered and still ongoing 9999 = unknown if administered, date unknown	Some bilateral cases may have different radiation treatments performed on both breasts on the same date. This file only captures radiation data on one of the breasts. Use the radiation dataset to get more detailed treatment info.
REG_MODALITY	Radiation treatment modality	See list of codes in Codebook.
RAD_BOOST_DOSE	Boost radiation dose cGy 8888 = radiation administered and still ongoing 9999 = unknown if administered, date unknown	
BOOST_MODALITY	Boost radiation treatment modality	Codes the same as REG_MODALITY

NBR_RXS_VOL	First course, total # of episodes of radiation (2-digit) Usually presented in the chart as (Reg + Boost) 99 = Unknown	If radiation was administered to both breasts, this data is from the tumor associated with FIRST_RAD_DATE.
RAD_SITES	Radiation sites	See list of codes in Codebook.
RAD_RX_CMPL_ST	First course, radiation completion status 0 = No radiation 1 = Radiation completed 2 = Radiation not completed, patient health 3 = Radiation not completed, patient expired 4 = Radiation not completed, patient choice 5 = Radiation not completed, family choice 6 = Radiation not completed, complications 7 = Radiation not completed, cytopenia 8 = Radiation not completed, other reasons 9 = Radiation not completed, unknown reasons	
RSN_NO_RAD	First course treatment, reason no radiation treatment 0 = Radiation done 1 = Rad not recomb; not planned as 1st course 2 = Contraindicated due to other causes 5 = Patient died prior to planned RT 6 = Recommended- Reason unknown no rad 7 = Patient/guardian refused 8 = Radiation recommended, unknown if done 9 = Unknown, death cert and autopsy only	

CHEMO_START_DATE	First chemotherapy start date (first course)	Obtained from the first primary tumor from the P3MA file.
CHEMO_CODE	<p>Chemo code</p> <p>00 = none, not part of 1st course rx</p> <p>01 = chemo as 1st course rx, type and number of agents not documented</p> <p>02 = chemo, single agent as 1st course tx</p> <p>03 = chemo, multi agents as 1st course tx</p> <p>82 = Chemo not recommended, contraindicated due to risk factors</p> <p>85 = Chemo not done, patient expired</p> <p>86 = chemo not done, was recommended, not stated</p> <p>87 = chemo not done, was recommended, patient/family member refused</p> <p>88 = chemo recomm, unk if given</p> <p>99 = unknown/ death certificate only</p>	
RSN_NO_CHEMO	<p>Reason no chemotherapy (first course)</p> <p>0 = Chemo done</p> <p>1 = Chemo not recommended</p> <p>2 = Contraindicated due to other causes, autopsy only cases</p> <p>6 = Reason unknown</p> <p>7 = Patient refused</p> <p>8 = Recommended, unknown if done</p> <p>9 = Unknown, death cert only cases</p>	

ADJCHEMO	Adjuvant chemotherapy, first course 0 = No 1 = Yes	Define as first course chemotherapy after the first surgery: CHEMO_DATE > FIRST_SURG_DATE
ADJCHEMO_START_DATE	Adjuvant chemotherapy start date, first course	
NEOCHEMO	Neoadjuvant chemotherapy, first course 0 = No 1 = Yes	Define as first course chemotherapy before the first surgery: CHEMO_DATE < FIRST_SURG_DATE
NEOCHEMO_START_DATE	Neoadjuvant chemotherapy start date, first course	
OTHERCHEMO	Other chemotherapy, first course 0 = No 1 = Yes	Define as first course chemotherapy in the absence of first course surgery.
OTHERCHEMO_START_DATE	Other chemotherapy start date, first course	

HORM_START_DATE	First hormone start date (first course)	Obtained from the first primary tumor from the P3MA file.
HORM_CODE	Hormone code (first course) 00 = none, not part of 1st course rx 01 = Hormone therapy administered as 1st course rx 82 = Hormone therapy not recommended, contraindicated due to risk factors 85 = Hormone therapy not done, patient expired 87 = Hormone therapy not done, was recommended, patient/family member refused 88 = Hormone therapy recommended, unk if done 99 = unknown/ death certificate only	
RSN_NO_HORM	Reason no hormone (first course) 0 = Hormone done 1 = Hormone not recommended 2 = Contraindicated due to other causes; autopsy only cases 6 = Reason unknown 7 = Patient/guardian refused 8 = Recommended, unknown if done 9 = Unknown, death certificate only cases	
LAST_HORM_END_DATE	Last hormone drug end date, first course	Missing if still taking drug. For women with multiple hormone drugs, this is the end date of the last hormone drug.

LAST_HORM_CMPL_ST	Last hormone drug completion status, first course 1 = Hormone treatment completed 2 = Not completed, patient health 3 = Not completed, patient expired 4 = Not completed, patient choice 5 = Not completed, family choice 6 = Not completed, complications 7 = Not completed, cytopenia 8 = Not completed, other reasons 9 = Not completed, unknown reason 10 = In progress	For women with multiple hormone drugs, this is the completion status of the last hormone drug.
TAMOXIFEN	Received Tamoxifen/Nolvadex, first course 0 = No 1 = Yes	Derived from Hormone file, agent code = 180973.
RALOXIFENE	Received Raloxifene/Evistan, first course 0 = No 1 = Yes	Derived from Hormone file, agent code = 706725.
ARIMIDEX	Received Arimidex/Anastrozole, first course 0 = No 1 = Yes	Derived from Hormone file, agent code = 719344.
AROMASIN	Received Aromasin/Exemestane, first course 0 = No 1 = Yes	Derived from Hormone file, agent code = 713563.
FEMARA	Received Femara/Letrozole, first course 0 = No 1 = Yes	Derived from Hormone file, agent code = 719345.

BIO_START_DATE	First biologics start date (first course)	Obtained from the first primary tumor from the P3MA file.
BIO_CODE	<p>First course therapy, biologics code</p> <p>00 = none, not part of 1st course rx</p> <p>01 = Biologics administered as 1st course rx</p> <p>82 = Biologics not recommended, contraindicated due to risk factors</p> <p>85 = Biologics not done, patient expired</p> <p>86 = Biologics not done, was recommended, not stated</p> <p>87 = Biologics not done, was recommended, patient/family member refused</p> <p>88 = Biologics recommended, unknown if done</p> <p>99 = unknown/ death certificate only</p>	
HERCEPTIN	<p>Received Herceptin/Trastuzumab, first course</p> <p>0 = No</p> <p>1 = Yes</p>	Derived from Biologics file, agent code = 688097.
HERCEPTIN_CMPL_ST	<p>Herceptin completion status, first course</p> <p>1 = completed</p> <p>2 = Not completed, patient health</p> <p>3 = Not completed, patient expired</p> <p>4 = Not completed, patient choice</p> <p>5 = Not completed, family choice</p> <p>6 = Not completed, complications</p> <p>7 = Not completed, cytopenia</p> <p>8 = Not completed, other reasons</p> <p>9 = Not completed, unknown reason</p> <p>10 = In progress</p>	For women with multiple Herceptin use, this is the completion status of the last treatment.

OTH_RX_START_DATE	First other treatment start date (first course)	Obtained from the first primary tumor from the P3MA file.
OTH_RX_CODE	First course, other treatment code 0 = No other cancer-directed therapy 1 = other CA dir treatment/bisphosphonates including SWOG 0307 trial 2 = other experimental rx 3 = double-blind study, code not broken 6 = unproved rx 7 = patient refused 8 = recomm, unkown if given 9 = unknown	
RSN_NO_OTHER	First course, reason no other treatment 0 = Other treatment done 1 = Not recommended 2 = Contraindicated 6 = Unknown 7 = Patient/guardian refused 8 = Recommended, unknown if done 9 = Death certificate cases only	

The followings is a set of treatment duration variables. The definition of treatment is different from previously defined variables.

Only first course surgery (within 18 months of diagnosis), radiation, and chemotherapy initiated before recurrence are counted as treatments for this set of variables. Hormone, biologics, and other treatments are excluded from consideration.

- 1) First cancer surgery – defined as the first surgery before recurrence and within 18 months of diagnosis.
- 2) Definitive surgery - defined as the last surgery before recurrence, within 18 months of diagnosis AND within 7 months of first surgery.

T_SURG1	Time to first treatment oriented surgery (days)	Time to first treatment-oriented surgery (SURG 20-90 and/or SLNS 3-7) from diagnosis and before recurrence. Surgery within 18 months of diagnosis.
T_DSURG	Time to final treatment oriented surgery (days)	Time to last treatment-oriented surgery from diagnosis <ul style="list-style-type: none"> • SURG 20-90 and/or SLNS 3-7 before recurrence date • Within 18 months of diagnosis AND within 7 months of first surgery.

SURGERY_TYPE2	Surgery type 1=Mastectomy 2=BCS 3=No surgery	Different from SURGERY_TYPE. This version only include surgeries within 18 months of diagnosis. <ul style="list-style-type: none"> • If patient has SURG values 30-80 in any row designate patient as “mastectomy” • If patient has only rows listing breast conserving surgery (code 20-29) and/or lymph node surgeries (code 3-7), designate as “BCS”. • Patients may also have rows indicating their biopsies (codes <20) but these don’t count as either one. •
T_CHEMO1	Time to first chemotherapy (days)	Calculated as time between diagnosis date and date of first chemotherapy before recurrence.
T_CHEMOL	Time to last chemotherapy (days)	Calculated as time between diagnosis date and date of last chemotherapy (D_CHEML). Note: chemo could end after recurrence as long as it is initiated before recurrence.
DUR_CHEMO	Duration of chemotherapy (days)	$T_CHEMOL - T_CHEMO1$
T_RAD1	Time to first radiation (days)	Calculated as time between diagnosis date and date of first radiation before recurrence.
T_RADL	Time to last radiation (days)	Calculated as time between diagnosis date and date of last radiation.
DUR_RAD	Duration of radiation (days)	$T_RADL - T_RAD1$

T_TREAT1	Time to first treatment (days)	The shortest of the following times: T_DSURG (definitive surgery), T_CHEM1, and T_RAD1.
T_TREATL	Time to last treatment (days)	The longest of the following times: T_DSURG (definitive surgery), T_CHEML, and T_RADL.
DUR_TREAT	Time to complete treatment	T_TREATL – T_TREAT1
T1_TYPE	First treatment type 1 = Surgery 2 = Radiation 3 = Chemotherapy 4 = Radiation & Chemotherapy	Derived from the 1 st surgery (T_SURG1), the 1 st chemo and the 1 st radiation. Note: first cancer directed surgery is used here instead of definitive cancer surgery.
PT_GROUP	Patient groups 1 = Surgery only 2 = Surgery + radiation 3 = Surgery + chemo 4 = Surgery + chemo + radiation	Undefined for cases without first course surgery within 18 months of diagnosis. 1) Surgery only: definitive surgery is the only treatment prior to recurrence 2) Surgery plus radiation: has a definitive surgery and a first radiotherapy prior to recurrence. This patient has no chemotherapy prior to recurrence. 3) Surgery plus chemo: has a definitive surgery and chemotherapy prior to recurrence, but no radiotherapy. 4) Surgery plus chemo and radiation: has a definitive surgery and chemotherapy and radiation prior to recurrence. Note: definitive surgery is used here.

DIABETES	Comorbidity – Diabetes 0 = No 1 = Yes	See list of ICD9 codes in Codebook.
COPD	Comorbidity – COPD and bronchiectasis 0 = No 1 = Yes	See list of ICD9 codes in Codebook.
HEART_DZ	Comorbidity –Heart Disease and Congestive Heart Failure 0 = No 1 = Yes	See list of ICD9 codes in Codebook.
OBESITY	Comorbidity – Obesity 0 = No 1 = Yes	See list of ICD9 codes in Codebook.
HYPERTENSION	Comorbidity – Hypertension 0 = No 1 = Yes	See list of ICD9 codes in Codebook.

RECURRENT	Ever had subsequent recurrent breast cancer 1 = Yes 0 = No	
RECUR_DATE1	First subsequent recurrence date	mm/dd/yyyy
RECUR_TYPE1	First subsequent recurrence type	See list of codes in Codebook (page 23).
RECUR_LRD1	Recurrence type for the first recurrence 1 = Local 2 = Regional 3 = Distant	Local: RECUR_TYPE1 = 04, 06, 10, 13, 14, 15, 16, 17 Regional: RECUR_TYPE1 = 20, 21, 22, 25, 26, 27, 30, 36 Distant: RECUR_TYPE1 = 40, 46, 51-62
RECUR1_MET1 – RECUR1_MET3	First recurrent metastasis (up to 3 variables) 0 = None 1 = Peritoneum 2 = Lung (limited) 3 = Pleura 4 = Liver 5 = Bone 6 = CNS 7 = Skin 8 = Lymph nodes (distant) 9 = Other	
RECUR_DATE2	Second subsequent recurrence date	mm/dd/yyyy
RECUR_TYPE2	Second subsequent recurrence type	See list of codes in Codebook (page 23).

RECUR_LRD2	Recurrence type for the second recurrence 1 = Local 2 = Regional 3 = Distant	Local: RECUR_TYPE2 = 04, 06, 10, 13, 14, 15, 16, 17 Regional: RECUR_TYPE2 = 20, 21, 22, 25, 26, 27, 30, 36 Distant: RECUR_TYPE2 = 40, 46, 51-62
RECUR2_MET1 – RECUR2_MET3	Second recurrent metastasis (up to 3 variables) 0 = None 1 = Peritoneum 2 = Lung (limited) 3 = Pleura 4 = Liver 5 = Bone 6 = CNS 7 = Skin 8 = Lymph nodes (distant) 9 = Other	
RECUR_DATE3	Third subsequent recurrence date	mm/dd/yyyy
RECUR_TYPE3	Third subsequent recurrence type	See list of codes in Codebook (page 23).
RECUR_LRD3	Recurrence type for the third recurrence 1 = Local 2 = Regional 3 = Distant	Local: RECUR_TYPE3 = 04, 06, 10, 13, 14, 15, 16, 17 Regional: RECUR_TYPE3 = 20, 21, 22, 25, 26, 27, 30, 36 Distant: RECUR_TYPE3 = 40, 46, 51-62

RECUR3_MET1 – RECUR3_MET3	Third recurrent metastasis (up to 3 variables) 0 = None 1 = Peritoneum 2 = Lung (limited) 3 = Pleura 4 = Liver 5 = Bone 6 = CNS 7 = Skin 8 = Lymph nodes (distant) 9 = Other	
RECUR_ANY_LRD	Recurrence type (any recurrence) 1 = Local 2 = Regional 3 = Distant	Derived from RECUR_LRD1 – RECUR_LRD3 using this Hierarchy: distant > regional > local Distant = if any recurrence is distant Regional = if no distant recurrence and any recurrence is regional Local = only local recurrence

CODEBOOK

SURG (surgery code for the first surgery)

- 01 Biopsy(inc, need, asp) to site other than primary
- 02 Biopsy(inc, need, asp) to primary site
- 19 Local tumor destruction, NOS
- 20 Partial mastectomy, NOS; less than total mastectomy, NOS
- 21 Partial mastectomy with nipple resection
- 22 Lumpectomy or excisional biopsy
- 23 Reexcision of the biopsy site for gross or microscopic residual disease
- 24 Segmental mastectomy (including wedge resection, quadrantectomy, tylectomy)
- 30 Subcutaneous mastectomy
- 40 Total (simple) mastectomy
- 41 Total mastectomy without removal of uninvolved contralateral breast
- 42 Total Mastectomy with removal of uninvolved contralateral breast
- 43 Total Mastectomy with Reconstruction NOS without removal of uninvolved contralateral breast
- 44 Total Mastectomy with Reconstruction with tissue without removal of uninvolved contralateral breast
- 45 Total Mastectomy with Reconstruction with implant without removal of uninvolved contralateral breast
- 46 Total Mastectomy with Reconstruction combined (with tissue and implant) without removal of uninvolved contralateral breast
- 47 Total Mastectomy Reconstruction NOS with removal of uninvolved contralateral breast
- 48 Total Mastectomy Reconstruction with tissue with removal of uninvolved contralateral breast
- 49 Total Mastectomy Reconstruction with implant with removal of uninvolved contralateral breast
- 50 Modified radical mastectomy
- 51 Modified radical mastectomy without removal of uninvolved contralateral breast
- 52 Modified radical mastectomy with removal of uninvolved contralateral breast
- 53 Modified radical mastectomy Reconstruction NOS without removal of uninvolved contralateral breast
- 54 Modified radical mastectomy Reconstruction w/ tissue without removal of uninvolved contralateral breast
- 55 Modified radical mastectomy Reconstruction w/ implant without removal of uninvolved contralateral breast
- 56 Modified radical mastectomy Reconstruction combined (tissue and implant) without removal of uninvolved contralateral breast
- 57 Modified radical mastectomy Reconstruction, NOS with removal of uninvolved contralateral breast
- 58 Modified radical mastectomy Reconstruction, with tissue with removal of uninvolved contralateral breast
- 59 Modified radical mastectomy Reconstruction, with implant with removal of uninvolved contralateral breast

- 60 Radical mastectomy
- 61 Radical mastectomy without removal of uninvolved contralateral breast
- 62 Radical mastectomy with removal of uninvolved contralateral breast
- 63 Modified radical mastectomy Reconstruction, combined (tissue and implant) with removal of uninvolved contralateral breast
- 64 Radical mastectomy Reconstruction NOS without removal of uninvolved contralateral breast
- 65 Radical mastectomy Reconstruction w/ tissue without removal of uninvolved contralateral breast
- 66 Radical mastectomy Reconstruction w/ implant without removal of uninvolved contralateral breast
- 67 Radical mastectomy Reconstruction combined (tissue and implant) without removal of uninvolved contralateral breast
- 68 Radical mastectomy Reconstruction, NOS with removal of uninvolved contralateral breast
- 69 Radical mastectomy Reconstruction, with tissue with removal of uninvolved contralateral breast
- 70 Extended radical mastectomy
- 71 Extended radical mastectomy without removal of uninvolved contralateral breast
- 72 Extended radical mastectomy with removal of uninvolved contralateral breast
- 73 Radical mastectomy Reconstruction, with implant with removal of uninvolved contralateral breast
- 74 Radical mastectomy Reconstruction, combined (tissue and implant) with removal of uninvolved contralateral breast
- 75 Total Mastectomy Reconstruction combined (tissue and implant) with removal of uninvolved contralateral breast
- 80 Mastectomy, NOS
- 90 Surgery, NOS
- 99 Unknown if surgery was performed, death certificate only

REG_MODALITY (Radiation treatment modality); BOOST_MODALITY (Boost radiation treatment modality)

- 00 No radiation
- 20 External Beam, NOS
- 21 Orthovoltage
- 22 Cobalt 60, Cesium 137
- 23 Photons (2-5MV)
- 24 Photons (6-10MV)
- 25 Photons (11-19MV)
- 26 Photons (>19 MV)
- 27 Photons (mixed energies)
- 28 Electrons
- 29 Photons and Electrons mixed
- 30 Neutrons
- 31 IMRT
- 32 Conformal or 3-D therapy
- 40 Photons
- 41 Stereotactic radiosurg
- 42 Linac Radiosurgery
- 43 Gamma Knife
- 50 Brachytherapy, NOS
- 51 Brachytherapy, intracavitary, LDR
- 52 Brachytherapy, intracavitary, HDL
- 53 Brachytherapy Interstitial, LDR
- 54 Brachytherapy, interstitial HDL
- 55 Radium
- 60 Radioisotopes, NOS
- 61 Strontium- 89
- 62 Strontium- 90
- 80 Combination Modality, specified
- 85 Combination Modality, NOS
- 98 Other, NOS
- 99 Unknown

RAD_SITES (Radiation sites)

00 No radiation treatment	22 Mantle, Mini-mantle
01 Eye/orbit	23 Lower extended field
02 Pituitary	24 Spine
03 Brain (NOS)	25 Skull
04 Brain (limited)	26 Ribs
05 Head and Neck (NOS)	27 Hip
06 Head and Neck (limited)	28 Pelvic bones
07 Glottis	29 Pelvis (NOS)
08 Sinuses	30 Skin
09 Parotid	31 Soft tissue
10 Chest/lung NOS	32 Hemibody
11 Lung (limited)	33 Whole body
12 Esophagus	34 Bladder and pelvis
13 Stomach	35 Prostate and pelvis
14 Liver	36 Uterus and cervix
15 Pancreas	37 Shoulder
16 Kidney	38 Extremity bone, NOS
17 Abdomen (NOS)	39 Inverted Y
18 Breast	40 Spinal cord
19 Breast/lymph nodes	41 Prostate
20 Chest wall	50 Thyroid
21 Chest wall/lymph nodes	60 Lymph node region, NOS
	98 Other
	99 Unknown

Comorbidity Groups

Variable	Relevant ICD9 Codes
Diabetes	24900 25000 25001 7902 79021 79022 79029 7915 7916 V4585 V5391 V6546 24901 24910 24911 24920 24921 24930 24931 24940 24941 24950 24951 24960 24961 24970 24971 24980 24981 24990 24991 2500 25002 25003 25010 25011 25012 25013 25020 25021 25022 25023 25030 25031 25032 25033 25040 25041 25042 25043 25050 25051 25052 25053 2506 25060 25061 25062 25063 2507 25070 25071 25072 25073 2508 25080 25081 25082 25083 2509 25090 25091 25092 25093
COPD and bronchiectasis	490 4910 4911 4912 49120 49121 49122 4918 4919 4920 4928 494 4940 4941 496
Heart Disease and Congestive Heart Failure	4100 41000 41001 41002 4101 41010 41011 41012 4102 41020 41021 41022 4103 41030 41031 41032 4104 41040 41041 41042 4105 41050 41051 41052 4106 41060 41061 41062 4107 41070 41071 41072 4108 41080 41081 41082 4109 41090 41091 41092 4110 4111 4118 41181 41189 412 4130 4131 4139 4140 41400 41401 41406 4142 4143 4144 4148 4149 V4581 V4582 78650 78651 78659 4150 4151 41512 41513 41519 4160 4161 4162 4168 4169 4170 4171 4178 4179 41410 41411 41412 41419 4291 4292 4293 4295 4296 42971 42979 42981 42982 42983 42989 4299 39891 4280 4281 42820 42821 42822 42823 42830 42831 42832 42833 42840 42841 42842 42843 4289
Obesity	2780 27800 27801 27802
Hypertension	4011 4019 4010 40200 40201 40210 40211 4029 40290 40291 4030 40300 40301 4031 40310 40311 4039 40390 40391 4040 40400 40401 40402 40403 4041 40410 40411 40412 40413 4049 40490 40491 40492 40493 40501 40509 40511 40519 40591 40599 4372

RECUR_TYPE (Subsequence recurrence type)

- 00 = None
- 04 = In situ recurrence of invasive tumor
- 06 = In situ recurrence of in situ tumor
- 10 = Local recurrence, NOS
- 13 = Local recurrence of invasive tumor
- 14 = Trocar recurrence of invasive tumor
- 15 = Both local and trocar recurrence of invasive tumor
- 16 = Local recurrence of in situ tumor, NOS
- 17 = Both local and trocar recurrence of in situ tumor
- 20 = Regional recurrence, insufficient information
- 21 = Recurrence of invasive tumor in adjacent tissue or organ(s) only
- 22 = Recurrence of invasive tumor in regional lymph nodes only
- 25 = Recurrence of invasive tumor in adjacent tissue or organ(s) and in regional lymph nodes
- 26 = Regional recurrence of an in situ tumor, NOS
- 27 = Recurrence of an in situ tumor in adjacent tissue or organ(s) and in regional lymph nodes at the same time
- 30 = Both regional recurrence of an invasive tumor in adjacent tissue or organ(s) and/or regional lymph nodes and local and/or trocar recurrence
- 36 = Both regional recurrence of an in situ tumor in adjacent tissue or organ(s) and/or regional lymph nodes and local and/or trocar recurrence
- 40 = Distant recurrence, insufficient information
- 46 = Distant recurrence of in situ tumor
- 51 = Distant recurrence of an invasive tumor in peritoneum only
- 52 = Distant recurrence of an invasive tumor in the lung only
- 53 = Distant recurrence of an invasive tumor in pleura only
- 54 = Distant recurrence of an invasive tumor in liver only
- 55 = Distant recurrence of an invasive tumor in bone only
- 56 = Distant recurrence of an invasive tumor in CNS only
- 57 = Distant recurrence of an invasive tumor in the skin only
- 58 = Distant recurrence of an invasive tumor in lymph node only
- 59 = Distant systemic recurrence of invasive tumor only
- 60 = Distant recurrence of an invasive tumor in single distant site and local, trocar and/or regional recurrence
- 62 = Distant recurrence of an invasive tumor in multiple sites
- 70 = Since diagnosis, patient has never been disease free
- 88 = Disease has recurred, but type of recurrence is unknown
- 99 = Unknown whether the disease recurred or if the patient was ever disease free