

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

Dataset: MA\_VARS\_11072025.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 ( <b>invasive or in situ</b> ) within six months of the date of diagnosis of the first invasive tumor.

BILATERAL_GP	Bilateral cases group 1 = Invasive in one breast and CIS in the other breast 2 = Invasive tumors in both breast	Only available for 67 bilateral cases (LAT=3), missing data for others. Note: Six cases with ipsilateral invasive tumors have missing data.
SIZE	Tumor size (mm) 998 =Inflammatory carcinoma; diffuse, widespread, ¾ or more of breast 990= microinvasion, microscopic focus of invasive tumor, invasive tumor is less than or equal to 1 mm 999 = unknown	Obtained from P3MA (ERS) file. Can also record size of inflammatory carcinoma if available.
ESTSIZE	Tumor size 1 = ≤ 2 cm 2 = >2 – 5 cm 3 = >5 cm	Recoded from SIZE. SIZE=998 classified as “>5 cm” per Melissa.

AJCC_GRP	<p>AJCC Stage</p> <p>1 = Stage I</p> <p>1A = Stage IA</p> <p>1B = Stage IB</p> <p>2A = Stage IIA</p> <p>2B = Stage IIB</p> <p>3A = Stage IIIA</p> <p>3B = Stage IIIB</p> <p>3C = Stage IIIC</p> <p>4 = Stage IV</p> <p>88 = Not applicable</p> <p>99 = Unknown</p>	<p>Obtained from P3MA (ERS) file.</p> <p>1 = Stage I (diagnosed before 1/1/2010)</p> <p>1A = Stage IA (diagnosed 2010 and beyond)</p> <p>1B = Stage IB (diagnosed 2010 and beyond)</p>
STAGE	<p>AJCC Stage</p> <p>1 = Stage I</p> <p>2 = Stage II</p> <p>3 = Stage III</p> <p>4 = Stage IV</p>	<p>Recoded from AJCC_GRP</p> <p>1 = 1, 1A, 1B</p> <p>2 = 2A, 2B</p> <p>3 = 3A, 3B, 3C</p> <p>4 = 4</p>
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>

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 &amp; 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 following:</p> <ol style="list-style-type: none"> <li>1) Number of nodes positive for malignancy &gt;0</li> <li>2) Staging - Lymph node metastasis</li> </ol> <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>Tumor grade</p> <p>1 = Well differentiated</p> <p>2 = Moderately differentiated</p> <p>3 = Poorly differentiated</p> <p>4 = Undifferentiated/Anaplastic 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>
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>	

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  99= Unknown</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.
DAYS_RX	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 DAYS_RX. 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)	Surgery is defined as: SURG (code 20-90) and/or SLNS (code 3-7)  For bilateral cases, if surgery dates are different, take the earliest date.
SURGERY_YN	Received surgery (first course) 0 = No 1 = Yes	Defined from FIRST_SURG_DATE. Missing surgery date is counted as NO surgery.
SURGERY_TYPE	Type of surgery (first course) 1 = Mastectomy 2 = Breast conserving surgery 3 = No surgery	Derived from <b>all first course</b> surgeries. If a case has multiple breast tumors, derived variable from all tumors.  Mastectomy: if subject has any first course SURG=30-80.  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.  No surgery: if subject has no surgery data or SURG=0-19 and has no SURG>19.

RSN_NO_SURG	Reason no surgery (first course) 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	
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_YN	Received radiation (first course) 0 = No 1 = Yes	Defined from FIRST_RAD_DATE.

RSN_NO_RAD	<p>First course treatment, reason no radiation treatment</p> <p>0 = Radiation done</p> <p>1 = Rad not recomm; not planned as 1st course</p> <p>2 = Contraindicated due to other causes</p> <p>5 = Patient died prior to planned RT</p> <p>6 = Recommended- Reason unknown no rad</p> <p>7 = Patient/guardian refused</p> <p>8 = Radiation recommended, unknown if done</p> <p>9 = Unknown, death cert and autopsy only</p>	
CHEMO_START_DATE	First chemotherapy start date (first course)	Obtained from the first primary tumor from the P3MA file.
CHEMO_YN	<p>Received chemotherapy (first course)</p> <p>0 = No</p> <p>1 = Yes</p>	Defined from CHEMO_START_DATE. Missing chemo date is counted as NO chemotherapy.
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><del>82 = Chemo not recommended, contraindicated due to risk factors</del></p> <p><del>85 = Chemo not done, patient expired</del></p> <p><del>86 = chemo not done, was recommended, not stated</del></p> <p><del>87 = chemo not done, was recommended, patient/family member refused</del></p> <p><del>88 = chemo recomm, unk if given</del></p> <p><del>99 = unknown/ death certificate only</del></p>	Codes 82-99 not used.

RSN_NO_CHEMO	Reason no chemotherapy (first course) 0 = Chemo done 1 = Chemo not recommended 2 = Contraindicated due to other causes, autopsy only cases 6 = Reason unknown 7 = Patient refused 8 = Recommended, unknown if done 9 = Unknown, death cert only cases	
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_YN	Received hormone therapy (first course) 0 = No 1 = Yes	Defined from HORM_START_DATE. Missing hormone date is counted as NO hormone therapy.
HORM_CODE	Hormone code (first course) 00 = none, not part of 1st course rx 01 = Hormone therapy administered as 1st course rx <del>82 = Hormone therapy not recommended, —— contraindicated due to risk factors</del> <del>85 = Hormone therapy not done, patient expired</del> <del>87 = Hormone therapy not done, was recommended, —— patient/family member refused</del> <del>88 = Hormone therapy recommended, unk if done</del> <del>99 = unknown/ death certificate only</del>	Codes 82-99 not used.
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_YN	Received biologics (first course) 0 = No 1 = Yes	Defined from BIO_START_DATE. Missing biologics date is counted as NO biologics treatment.
BIO_CODE	First course therapy, biologics code 00 = none, not part of 1st course rx 01 = Biologics administered as 1st course rx <del>82 = Biologics not recommended, contraindicated — due to risk factors</del> <del>85 = Biologics not done, patient expired</del> <del>86 = Biologics not done, was recommended, not stated</del> <del>87 = Biologics not done, was recommended, — patient/family member refused</del> <del>88 = Biologics recommended, unknown if done</del> 99 = unknown/ death certificate only	Codes 82-99 not used.
HERCEPTIN	Received Herceptin/Trastuzumab, first course 0 = No 1 = Yes	Derived from Biologics file, agent code = 688097.
HERCEPTIN_CMPL_ST	Herceptin completion status, first course 1 = 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 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_YN	Received other treatment (first course) 0 = No 1 = Yes	Defined from OTH_RX_START_DATE. Missing date is counted as NO other treatment.
OTH_RX_CODE	First course, other treatment code 0 = No other cancer-directed therapy 1 = other CA dir treatment/bisphosphonates including SWOG 0307 trial <del>2 = other experimental rx</del> <del>3 = double blind study, code not broken</del> 6 = unproved rx 7 = patient refused 8 = recomm, unkown if given 9 = unknown	Codes 2-9 not used.
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 following 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 (SURG1) – defined as the first surgery before recurrence and within 18 months of diagnosis.**
- 2) Definitive surgery (DSURG) - 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"> <li>• SURG 20-90 and/or SLNS 3-7 before recurrence date</li> <li>• Within 18 months of diagnosis AND within 7 months of first surgery.</li> </ul>

SURGERY_TYPE2	Surgery type 1=Mastectomy 2=BCS 3=No surgery	Different from SURGERY_TYPE. <b>This version only includes surgeries within 18 months of diagnosis AND within 7 months of first surgery.</b> <ul style="list-style-type: none"> <li>• If patient has SURG values 30-80 in any row designate patient as “mastectomy”</li> <li>• If patient has only rows listing breast conserving surgery (code 20-29) and/or lymph node surgeries (code 3-7), designate as “BCS”.</li> <li>• Patients may also have rows indicating their biopsies (codes &lt;20) but these don’t count as either one.</li> </ul>
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 ( <b>definitive surgery</b> ), T_CHEMO1, and T_RAD1.
T_TREATL	Time to last treatment (days)	The longest of the following times: T_DSURG ( <b>definitive surgery</b> ), T_CHEMOL, 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 <sup>st</sup> surgery (T_SURG1), the 1 <sup>st</sup> chemo and the 1 <sup>st</sup> radiation.  <b>Note: first cancer directed surgery is used here instead of definitive cancer surgery.</b>
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: <b>definitive surgery</b> is the only treatment prior to recurrence 2) Surgery plus radiation: has a <b>definitive surgery</b> and a first radiotherapy prior to recurrence. This patient has no chemotherapy prior to recurrence. 3) Surgery plus chemo: has a <b>definitive surgery</b> and chemotherapy prior to recurrence, but no radiotherapy. 4) Surgery plus chemo and radiation: has a <b>definitive surgery</b> and chemotherapy and radiation prior to recurrence.  <b>Note: definitive surgery is used here.</b>

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 24).
RECUR_LRD1	Recurrence type for the first recurrence 1 = Local 2 = Regional 3 = Distant 9 = Unknown	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 Unknown: RECUR_TYPE1 = 88
RECUR_NODE1	Recurrence #1 – regional node status 1 = Positive 2 = Negative	Define for RECUR_LRD1=2 (regional) only. Missing for Local/Distant recurrence Positive: RECUR_TYPE1 = 20, 22, 25, 30 Negative: RECUR_TYPE1 = 21, 26, 27, 36
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 24).
RECUR_LRD2	Recurrence type for the second recurrence 1 = Local 2 = Regional 3 = Distant 9 = Unknown	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 Unknown: RECUR_TYPE2 = 88
RECUR_NODE2	Recurrence #2 – regional node status 1 = Positive 2 = Negative	Define for RECUR_LRD2=2 (regional) only. Missing for Local/Distant recurrence Positive: RECUR_TYPE2 = 20, 22, 25, 30 Negative: RECUR_TYPE2 = 21, 26, 27, 36
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 24).
RECUR_LRD3	Recurrence type for the third recurrence 1 = Local 2 = Regional 3 = Distant 9 = Unknown	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 Unknown: RECUR_TYPE2 = 88
RECUR_NODE3	Recurrence #3 – regional node status 1 = Positive 2 = Negative	Define for RECUR_LRD3=2 (regional) only. Missing for Local/Distant recurrence Positive: RECUR_TYPE3 = 20, 22, 25, 30 Negative: RECUR_TYPE3 = 21, 26, 27, 36
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	<p>Recurrence type (most severe recurrence)</p> <p>1 = Local 2 = Regional 3 = Distant 9 = Unknown</p>	<p>Derived from RECUR_LRD1 – RECUR_LRD3 using this Hierarchy: distant &gt; regional &gt; local</p> <p>Distant = if any recurrence is distant Regional = if no distant recurrence and any recurrence is regional Local = only local recurrence</p> <p>Note: A case may have recurrences with different types. This variable captures the most severe recurrence type.</p>
RECUR_ANY_NODE	<p>Recurrence – any regional node status</p> <p>1 = Positive 2 = Negative</p>	<p>Positive if any one of the 3 variables: RECUR_NODE1, RECUR_NODE2, RECUR_NODE3 is positive.</p>
DAYS_RECUR1	<p>Days from diagnosis to the first recurrence</p>	<p>For cases with recurrent = RECUR_DATE1 – DX_DATE</p> <p>For cases without recurrent = DT_LST_CONT – DX_DATE</p>

## CODEBOOK

## Comorbidity Groups

Variable	Relevant ICD9 Codes
<b>Diabetes</b>	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
<b>COPD and bronchiectasis</b>	490 4910 4911 4912 49120 49121 49122 4918 4919 4920 4928 494 4940 4941 496
<b>Heart Disease and Congestive Heart Failure</b>	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
<b>Obesity</b>	2780 27800 27801 27802
<b>Hypertension</b>	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