FRAMINGHAM HEART STUDY CODING MANUAL

Cancer

COHORT: Original, Offspring, Generation 3, New Offspring Spouse, Omni 1, Omni 2

TIME INTERVAL: Cancer Cases Reviewed through December 2010

DATASET NAME: vr_cancer_2013_a_0018d

OBSERVATIONS: 5476

Variables that define a unique record: PID, TOPO, LOCATION, HIST, BEHAVIOR, GRADE.

This dataset has 5476 cancer cases among 4242 individuals.

All known cancer cases were validated using medical records (pathology reports). Occasionally a pathology report was unavailable. In these cases, other confirming evidence was used (see DIAG_TYP variable). This dataset contains only primary cancers (ICD-O behavior codes of 2 or 3). This dataset does not include metastatic cancers (ICD-O behavior code of 6). Only first occurrence of site and cell type is included in this dataset.

Most cancer records were reviewed using the FHS version of the 1976 ICD-O-1 code book. Refer to ICD-O-3 for details of morphology codes of lymphomas and several other morphologies.

To exclude non-melanotic skin cancers from this dataset:
Delete cases where (TOPO = 173 and(HIST < 8720 or HIST => 8790)).

This dataset replaces vr_cancer_2011_m_0690.

This dataset is a modified version of the original dataset. Some observations were deleted due to consent restrictions. This coding manual is based on the original dataset. Therefore the value ranges, number of observations and/or cohorts found in the modified dataset, may differ from those stated in this manual.

IDTYPE

Variable name: IDTYPE
Description: Framingham Heart Study cohort identifier

Code or Value	Description	Count
0	Original Cohort	2446
1	Offspring Cohort	2512
2	New Offspring Spouse Cohort	42
3	Generation 3 Cohort	389
7	Omni 1 Cohort	66
72	Omni 2 Cohort	21

ID

Variable name: ID
Description: Framingham Heart Study participant identifier within a cohort

Note: IDTYPE and ID taken together form a unique study participant identification number.

ID deleted - use random ID (PID)

PID

Variable name: PID
Description: Random ID

DIAG_TYP

Variable name: DIAG_TYP
Description: Method of diagnosis - represents the basis of diagnosis

Code or Value	Description	Count
1	Microscopically confirmed	5304
2	Clinically diagnosed	165
3	Death certificate only	7

D_DATE

Variable name: D_DATE
Description: Diagnosis date of tumor or cancer
Units: number of days since exam 1

Code or Value	Description	Count
-10345 — 22506	Range of values	5476

Note: Date was taken from pathology reports if available. Date was taken from other records if pathology report was not available or if diagnosis was made before the biopsy then the variable EXACT = 1.

If month and year were available, but not day, the 15th was used as the day, and the variable EXACT = 1. If only the year was available, JULY 4 was used as the month and day, and the variable EXACT = 1.

In some cases if biopsy was done at autopsy, then D_DATE was changed to precede DATE OF DEATH by one day.

D_DATE includes month, day and year, however this coding manual displays only month and year to preserve confidentiality.

EXACT

Variable name: EXACT
Description: Type of Date - diagnosis date qualifier - indicates the quality of D_DATE

Code or Value	Description	Count
0	Month, date and year of diagnosis date taken from pathology report or other test report.	5147
1	Estimated diagnosis date as described in variable D_DATE	329

TOPO

Variable name: TOPO

Description: Topography - describes site - corresponds to first three digits of topography code WHO ICD-O

Code or Value	Description	Count
140	Lip	0
141	Tongue	13
142	Major salivary gland	6
143	Gum	0
144	Floor of mouth	9
145	Other and unspecified parts of mouth	15
146	Oropharynx	12
148	Hypopharynx	9
149	Pharynx, Nasopharynx and ill-defined sites in lip, oral cavity and pharynx	7
150	Esophagus	29
151	Stomach	101
152	Small intestine	14
153	Colon	379
154	Rectum, rectosigmoid junction, anal canal and anus, not otherwise specified	147
155	Liver and intrahepatic bile ducts	23
156	Gallbladder and extrahepatic bile ducts	25
157	Pancreas	90
158	Retroperitoneum and peritoneum	14
160	Nasal cavities, accessory sinuses, middle ear and inner ear	7
161	Larynx	39
162	Trachea bronchus, and lung	464
163	Pleura	10
169	Hematopoietic and reticuloendothelial systems	189
170	Bones, joints and articular cartilage	9
171	Connective, subcutaneous and other soft tissues	22
173	Skin (excludes skin of labia majora, skin of vulva, skin of penis and skin of scrotum)	1776

174	Female breast (excludes skin of breast)	668
175	Male breast (excludes skin of breast)	10
179	Uterus, not otherwise specified	5
180	Cervix uteri	72
181	Placenta	0
182	Corpus uteri or endometrium	100
183	Ovary, fallopian tube and broad ligament	65
184	Other and unspecified female genital organs	17
185	Prostate gland	489
186	Testis	19
187	Penis and other male genital organs	5
188	Urinary bladder	216
189	Kidney and other urinary organs	94
190	Eye and lacrimal gland	8
191	Brain	55
193	Thyroid gland	51
195	Other ill-defined sites	10
196	Lymph nodes	107
199	Unknown primary site	76

Note: TOPO modified - cancer site categories with less than five cases were recategorized under "other" to preserve participant confidentiality.

LOCATION

Variable name: LOCATION

Description: Location - describes subdivision of site - corresponds to fourth digit of topography code WHO ICD-O

Code or Value	Description	Count
0 — 9	Range of values	5476

HIST

Variable name: HIST

Description: Histology or morphology - describes cell histopathology - corresponds to four digit morphology code WHO ICD-O

Code or Value	Description	Count
8000 — 9990	Range of values	5476

Note: Additional morphology codes from the ICD-O-3 code book are added to the FHS cancer coding system:

8046 - Non-small cell carcinoma

8077 - Squamous intraepithelial neoplasia,
      Cervical intraepithelial neoplasia,
      Vaginal intraepithelial neoplasia,
      Vulvar intraepithelial neoplasia,
      Anal intraepithelial neoplasia

8130 - Papillary transitional cell carcinoma

8148 - Glandular intraepithelial neoplasia,
      Prostatic intraepithelial neoplasia

8162 - Klatskin tumor

8253 - Bronchiolo-alveolar carcinoma, mucinous

8255 - Adenocarcinoma with mixed subtypes

8317 - Renal cell carcinoma, chromophobe type

8522 - Intraductal carcinoma and lobular carcinoma in situ

8858 - Dedifferentiated liposarcoma

9590-9769 - Hodgkin's and non-Hodgkin lymphomas

9590-9596 - Malignant lymphomas, NOS or diffuse

9650-9667 - Hodgkin's lymphoma

9670-9729 - Non-Hodgkin's lymphomas

9670-9699 - Mature B-cell lymphomas

    9673 - Mantle cell lymphoma

    9680 - Malignant lymphoma, large B-cell, diffuse or malignant
        lymphoma, large cell

9700-9719 - Mature T-cell and NK-cell lymphomas

9727-9729 - Precursor cell lymohoblastic lymphomas

9731-9734 - Plasma cell tumors

9740-9742 - Mast cell tumors

9750-9758 - Neoplasms of histiocytes and accessory lymphoid cells

9760-9769 - Immunoproliferative diseases

9872 - Acute myeloid leukemia, minimal differentiation

9895 - Acute myeloid leukemia with multilineage dysplasia

9987 - Therapy related myelodysplastic syndrome

9989 - Myelodysplastic syndrome, NOS

BEHAVIOR

Variable name: BEHAVIOR

Description: Behavior - decribes the degree of malignancy - corresponds to the fifth digit behavior code WHO ICD-O

Code or Value	Description	Count
0	Benign	0
1	Uncertain, borderline	0
2	Carcinoma in situ, non-invasive	436
3	Malignant, primary site	5040
6	Malignant, metastatic site	0
9	Malignant, uncertain whether primary or metastatic site	0

Note: This set contains only cases where BEHAVIOR = 2 or BEHAVIOR = 3. Metastatic code (BEHAVIOR = 6) was not used; rather the location of the primary site was used. If the primary site is unknown, the code 199 is used for the site.

GRADE

Variable name: GRADE

Description: Grade - correspnds to code for histologic grading and differentiation WHO ICD-O

Code or Value	Description	Count
1	Grade I - well-differentiated, differentiated	513
2	Grade II - moderately well-differentiated, moderately differentiated	724
3	Grade III - poorly differentiated	886
4	Grade IV - undifferentiated, anaplastic	239
9	Grade or differentiation not determined, stated, or applicable	3114

ESTRGN

Variable name: ESTRGN
Description: Estrogen receptors

Code or Value	Description	Count
0	Negative	49
1	Weakly positive	10
2	Positive	240
8	Not done	136
.	Unknown or collected on a form that did not contain this variable	5041

Note: This variable was added to the data collection form in version 10-01-07.

PRGSTRN

Variable name: PRGSTRN
Description: Progesterone receptors

Code or Value	Description	Count
0	Negative	72
1	Weakly positive	22
2	Positive	192
8	Not done	147
.	Unknown or collected on a form that did not contain this variable	5043

Note: This variable was added to the data collection form in version 10-01-07.

H2NI

Variable name: H2NI

Description: HER2NEU (human epidermal growth factor receptor 2) by IHC (immunohistochemistry)

Code or Value	Description	Count
0	Negative, 0 or 1+	138
1	Weakly positive, 2+	11
2	Positive, 3+	20
8	Not done	333
.	Unknown or collected on a form that did not contain this variable	4974

Note: This variable was added to the data collection form in version 10-01-07.

H2NF

Variable name: H2NF

Description: HER2NEU (human epidermal growth factor receptor 2) by FISH (fluorescence in situ hybridization)

Code or Value	Description	Count
0	Unamplified	35
1	Amplified by FISH	5
8	Not done	462
.	Unknown or collected on a form that did not contain this variable	4974

Note: This variable was added to the data collection form in version 10-01-07.

TISSUE_1

Variable name: TISSUE_1

Description: Tissue #1 related to Gleason score for prostate cancer

Code or Value	Description	Count
1 — 5	Range of values	281
.	Unknown or collected on a form that did not contain this variable	5195

Note: Classifying prostate cancer cells in terms of how abnormal they appear when examined under a microscope. Grades range from 1 to 5 with 1 being the lease aggressive and 5 the most aggressive.

This variable was added to the data collection form in version 10-01-07.

TISSUE_2

Variable name: TISSUE_2

Description: Tissue #2 related to Gleason score for prostate cancer

Code or Value	Description	Count
1 — 5	Range of values	269
.	Unknown or collected on a form that did not contain this variable	5207

Note: Classifying prostate cancer cells in terms of how abnormal they appear when examined under a microscope. Grades range from 1 to 5 with 1 being the lease aggressive and 5 the most aggressive.

This variable was added to the data collection form in version 10-01-07.

GLEASON

Variable name: GLEASON

Description: Gleason score: classifying prostate cancer cells

Code or Value	Description	Count
2 — 10	Range of values	305
.	Unknown or collected on a form that did not contain this variable	5171

Note: Gleason score (total between 2-10) TISSUE_1 + TISSUE_2 = GLEASON
If only GLEASON is given, TISSUE_1 and TISSUE_2 = "."
If no scores are given, GLEASON = "." and TISSUE_1 and TISSUE_2 = "."

This variable was added to the data collection form in version 10-01-07.

VERSION

Variable name: VERSION

Description: Version date of the data collection form

Code or Value	Description	Count
10-01-2007	Data collection form revised on 10-01-2007	551
03-15-2012	Data collection form revised on 03-15-2012	204
.	Unknown	4721

Note: Prior to 10-01-2007 data collection forms did not have version dates

VERSION deleted - for administrative purposes only

INDICATOR

Variable name: INDICATOR

Description: Indicates ICD-O code book used for coding

Code or Value	Description	Count
0	ICD-O-1 code book	5288
1	ICD-O-3 code book	155
2	ICD-O-3 code book using 'Coding Guidelines for Topography and Morphology' in front of code book	33