

HIP SCREENING STUDY: DATA DICTIONARY

TABLE OF CONTENTS

Document Summary.....	2
HIP screening study: Data Dictionary	3
Section 1: Identifiers	3
Section 2: Time Dependant	4
Section 3: Screening	6
Section 4: Breast Cancer Demographics	8
Section 5: Breast Cancer Diagnosis	9
Section 6: Breast Cancer Characteristics	11
Section 7: Breast Cancer Treatment	13
Section 8: Breast Cancer Mortality	14
Section 9: Death	15

Document Summary

Property	Value
Document Title	HIP screening study: Data Dictionary
Date Created	07/26/2019
For Dataset	hip_data_d04152015
Sections	9
Entries	63
Document Filename	dictionary_hip.d041515.rtf

HIP screening study: Data Dictionary
07/26/2019
Dataset: hip_data_d04152015

HIP screening study: Data Dictionary

Section 1: Identifiers

Variable	Label	Description	Format Text
cohort	Study cohort		1="Study Screened" 2="Study Refused Screening" 3="Control"
iid	IID		Character, 7
pair_iid	Matched iid		Character, 7

Section 2: Time Dependant

Variable	Label	Description	Format Text
age_entry	Age at entry		Numeric
followup_days	Days from randomization to date last followup		Numeric .I="Invalid Data" .M="Missing Data"
revised_entry_fu1	Revised interval from entry to last followup 1		Numeric .U="Unknown"
revised_entry_fu2	Revised interval from entry to last followup 2		Numeric .U="Unknown"
revised_entry_fu3	Revised interval from entry to last followup 3		Numeric .U="Unknown"
revised_survey_stat05	Revised status at year 5		1="Alive and no Breast Cancer" 2="Dead, no Breast Cancer" 3="Breast Cancer Diagnosed" 4="Post Office Return" 5="Non-Respondent to Survey" 6="Reported Death" 9="Unknown"
revised_survey_stat10	Revised status at year 10		1="Alive and no Breast Cancer" 2="Dead, no Breast Cancer" 3="Breast Cancer Diagnosed" 4="Post Office Return" 5="Non-Respondent to Survey" 6="Reported Death" 9="Unknown"
revised_survey_stat15	Revised status at year 15		1="Alive and no Breast Cancer" 2="Dead, no Breast Cancer" 3="Breast Cancer Diagnosed" 4="Post Office Return" 5="Non-Respondent to Survey" 6="Reported Death" 9="Unknown"
survey_stat05	Survey status 5 year		1="Alive, no Breast Cancer" 2="Death at any Time" 3="Breast Cancer at any Time" 4="Post Office Return" 5="Non Respondent to Survey" 6="Reported Death"
survey_stat10	Survey status 10 year		1="Alive, no Breast Cancer" 2="Death at any Time" 3="Breast Cancer at any Time" 4="Post Office Return" 5="Non Respondent to Survey" 6="Reported Death"

Variable	Label	Description	Format Text
survey_stat15	Survey status 15 year		1="Alive, no Breast Cancer" 2="Death at any Time" 3="Breast Cancer at any Time" 4="Post Office Return" 5="Non Respondent to Survey" 6="Reported Death"

Section 3: Screening

Variable	Label	Description	Format Text
initial_screen_days	Days from randomization to date of initial screen		Numeric .I="Invalid Data" .N="Not Applicable"
recom1	First annual exam recommendations		0="Not a Study Screened Woman" 1="Routine One Year Exam" 2="Early Reexamination - Clinical + Radiologic Recommendation" 3="Early Reexamination - Clinical Only" 4="Early Reexamination - Radiologic Only" 5="Biopsy or Aspiration" 6="Other"
recom2	Second annual exam recommendations		0="Not a Study Screened Woman" 1="Routine One Year Exam" 2="Early Reexamination - Clinical + Radiologic Recommendation" 3="Early Reexamination - Clinical Only" 4="Early Reexamination - Radiologic Only" 5="Biopsy or Aspiration" 6="Other"
recom3	Third annual exam recommendations		0="Not a Study Screened Woman" 1="Routine One Year Exam" 2="Early Reexamination - Clinical + Radiologic Recommendation" 3="Early Reexamination - Clinical Only" 4="Early Reexamination - Radiologic Only" 5="Biopsy or Aspiration" 6="Other"
recomi	Initial exam recommendations		0="Not a Study Screened Woman" 1="Routine One Year Exam" 2="Early Reexamination - Clinical + Radiologic Recommendation" 3="Early Reexamination - Clinical Only" 4="Early Reexamination - Radiologic Only" 5="Biopsy or Aspiration" 6="Other"
scr_pat	Screening exam pattern		0="Not a Study Screened Woman" 1="Initial Exam Only" 2="Initial and First" 3="Initial and Second" 4="Initial and Third" 5="Initial, First and Second" 6="Initial, Second and Third" 7="Initial, First, and Third" 8="Initial, First, Second, and Third"
screen1_days	Days from randomization to date of first annual screen		Numeric .I="Invalid Data" .N="Not Applicable"
screen2_days	Days from randomization to date of second annual screen		Numeric .I="Invalid Data" .N="Not Applicable"

Variable	Label	Description	Format Text
screen3_days	Days from randomization to date of third annual screen		Numeric .!="Invalid Data" .N="Not Applicable"

Section 4: Breast Cancer Demographics

Variable	Label	Description	Format Text
marital	Marital status at dx		.C="Not a Breast Cancer Case" .V="Not Collected" 1="Single, Never Married" 2="Married" 3="Widowed" 4="Separated or Divorced" 9="Unknown"
menst	Menstrual status at dx		.C="Not a Breast Cancer Case" .V="Not Collected" 1="Still Menstruating, Periods Regular" 2="Still Menstruating, Periods Irregular" 3="Still Menstruating, Period Regularity Unknown" 4="Post-Menstrual, Cessation Natural" 5="Post-Menstrual, Cessation Surgical With Bilateral Oophorectomy" 6="Post-Menstrual, Cessation Surgical Without Oophorectomy" 7="Post-Menstrual, Cessation Surgical, Oophorectomy Status Unknown" 8="Post-Menstrual, Natural vs. Artificial Unknown" 9="Unknown"
menst_yrs	Years post-menstrual at dx		.C="Not a Breast Cancer Case" .V="Not Collected" 1="< 1 Year" 2="1.0-2.9 Years" 3="3.0-4.9 Years" 4="5.0-9.9 Years" 5="10 Years or More" 6="Still Menstruating" 9="Unknown"
race3	Race, reported at diagnosis		.C="Not a Breast Cancer Case" .V="Not Collected" 1="White" 2="Black" 3="Other" 9="Unknown"

Section 5: Breast Cancer Diagnosis

Variable	Label	Description	Format Text
age_diag	Age at diagnosis		Numeric .C="Not a Breast Cancer Case"
cancer_flag	Breast cancer flag		0="No Breast Cancer" 1="Breast Cancer, Interval From Entry to Diagnosis <= 15 Years" 2="Breast Cancer, Interval From Entry to Diagnosis > 15 Years" 9="Status Unknown"
case_detec	Detection classification and evidence for biopsy recommendation		.C="Not a Breast Cancer Case" 1="Clinical Only" 2="Radiologic Only, Mass On X-Ray" 3="Radiologic Only, Microcalcifications On X-Ray" 4="Radiologic Only, Mass + Microcalcifications" 5="Clinical + Radiologic, Both Mass" 6="Clinical + Radiologic, Clinical Mass And Radiologic Microcalcifications" 8="Interval Diagnosis Less Than 12 Months After Last Screening" 9="Interval Diagnosis 12 Months Or More After Last Screening" 10="Study Refused Screening" 12="Control"
case_lscr	Last screening exam prior to diagnosis		.C="Not a Breast Cancer Case" .U="Unknown" 0="Initial Screening" 1="Initial On Early Recall Exam" 10="First Annual Exam" 11="First Annual On Early Recall" 20="Second Annual Exam" 21="Second Annual On Early Recall" 30="Third Annual Exam" 31="Third Annual On Early Recall" 44="Last Exam Initial, Detection Not Attributed To Screening" 45="Last Exam First Annual, Detection Not Attributed To Screening" 46="Last Exam Second Annual, Detection Not Attributed To Screening" 47="Last Exam Third Annual, Detection Not Attributed To Screening" 49="No Histologic Confirmation Prior To Death" 99="Not A Screened Woman"
diag_days	Days from randomization to date of diagnosis		Numeric .N="Not Applicable"

Variable	Label	Description	Format Text
path_source	Pathology source		.C="Not a Breast Cancer Case" 0="Unknown" 1="Coordinating Pathologist" 2="Hospital Pathologist, Review of Report by Coordinating Pathologist" 3="Hospital Tumor Registry, No Central Review of Slides or Report" 4="New York State Cancer Registry Only, Unable to Access Slides or Report"

Section 6: Breast Cancer Characteristics

Variable	Label	Description	Format Text
brst_side	Breast operated, malignant		.C="Not a Breast Cancer Case" 1="Right" 2="Left" 3="Both (Simultaneous Diagnosis)" 9="Unknown"
hist	Histologic type		.C="Not a Breast Cancer Case" 0="Unknown" 1="Duct Cell Carcinoma" 2="Intraductal Comedo" 3="Intraductal Papillary" 4="Duct Cell + Intraductal, Mixed" 5="Colloid" 6="Medullary" 7="Lobular In-Situ" 8="Sarcoma or Cystosarcoma" 9="Infiltrating Lobular" 10="Carcinoma And Infiltrating Carcinoma Type Unspecified" 11="Unknown" 12="Intraductal, Type Unspecified" 13="Other"
lesion_site	Site of lesion		.C="Not a Breast Cancer Case" .V="Not Collected" 1="Upper Outer Quadrant" 2="Lower Outer" 3="Upper Inner" 4="Lower Inner" 5="Upper Middle" 6="Lower Middle" 7="Outer Lateral Middle" 8="Inner Medial Middle" 9="Central" 10="Axilla" 11="Unknown" 12="Any Combination 1-9" 13="Upper, Inner/Outer Not Specified" 14="Lower, Inner/Outer Not Specified" 15="Outer, Upper/Lower Not Specified" 16="Inner, Upper/Lower Not Specified"
lesion_size	Size of lesion per hospital report		.C="Not a Breast Cancer Case" .V="Not Collected" 1="2.0 cm or Less" 2="2.1-4.0 cm" 3="4.1 cm or More" 4="2.0 cm or Less With Residual in Mastectomy Specimen" 5="2.1-4.0 cm With Residual in Mastectomy Specimen" 6="4.1 cm or More With Residual in Mastectomy Specimen" 9="Unknown, Not Measured or Not Stated in Hospital Report"

Variable	Label	Description	Format Text
node_num	Number of axillary nodes isolated per hospital report		.C="Not a Breast Cancer Case" .V="Not Collected" 0="None" 1="One" 2="Two" 3="Three" 4="Four" 5="Five" 6="Six" 7="Seven or More" 8="Several or Many" 9="Unknown"
node_pos	Number of axillary nodes positive per hospital report		.C="Not a Breast Cancer Case" .V="Not Collected" 0="None" 1="One" 2="Two" 3="Three" 4="Four" 5="Five" 6="Six" 7="Seven or More" 8="Several or Many" 9="Unknown"
node_stat	Axillary lymph node status		.C="Not a Breast Cancer Case" 0="Unknown, No Information (No Nodes Identified Per Hospital Report or Unable to Access Hospital Records)" 1="Negative" 2="Positive" 3="Unknown, No Nodes Removed" 9="Unknown, No Information (No Nodes Identified Per Hospital Report or Unable to Access Hospital Records)"
stage	Extent of disease		.C="Not a Breast Cancer Case" .V="Not Collected" 1="Stage I" 2="Stage II" 3="Stage III" 4="Stage IV" 5="Clinical Stage I or II (No Axillary Nodes Removed For Staging, but no Demonstrable Evidence that Case Should Be Classified in Stages III or IV)" 9="Unknown"

Section 7: Breast Cancer Treatment

Variable	Label	Description	Format Text
radiation	Local radiation therapy		.C="Not a Breast Cancer Case" .V="Not Collected" 0="Unknown" 1="None" 2="Postop Cobalt" 3="Postop 200-250 KV" 4="Postop, Other" 5="Postop, Type Unknown" 6="Preop Cobalt" 7="Preop 200-250 KV" 8="Preop, Other" 9="Preop, Type Unknown"
surgery	Type of surgery		.C="Not a Breast Cancer Case" 1="Biopsy (Incisional, Needle) Only" 2="Local Excision Or Partial Mastectomy" 3="Total (Simple) Mastectomy" 4="Total (Simple) Mastectomy With Axillary Node Dissection" 5="Modified Radical Mastectomy" 6="Radical Mastectomy" 7="Extended Radical Mastectomy" 9="Unknown" 10="Biopsy of Metastatic Lesion, Primary Established As Breast" 11="Biopsy or Excision of Axillary Node, Primary Breast" 12="None Of The Above" 13="Local Excision Or Partial Mastecomy With Axillary Node Dissection"

Section 8: Breast Cancer Mortality

Variable	Label	Description	Format Text
age_exit	Age at death, or last followup if alive		Numeric .C="Not a Breast Cancer Case"
case_rvwd_x	Case reviewed by reviewer x		.C="Not a Breast Cancer Case" .I="Invalid Data" 0="No" 1="Yes"
case_rvwd_y	Case reviewed by reviewer y		.C="Not a Breast Cancer Case" 0="No" 1="Yes"
classif	Classification for data analysis (after death review)		.C="Not a Breast Cancer Case" 1="Both Life Table And Mortality Data (Deaths From Breast Cancer Among Histologically Confirmed Cases)" 2="Life Table Only (Living Women Plus Deaths Due To Other Than Breast Cancer Among Histologically Confirmed Cases)" 3="Mortality Data Only (Cases Not Histologically Confirmed Prior To Death From Breast Cancer)"
final_classification	Final classification		.C="Not a Breast Cancer Case" 0="Not Reviewed, or No Difference of Opinion on Underlying Cause of Death" 1="Cause Breast Cancer" 2="Cause Other Than Breast Cancer"
mort_class_x	Mortality classification, reviewer x		.C="Not a Breast Cancer Case" 0="Not Reviewed" 1="Reasonably Certain Underlying Cause of Death is Breast Cancer" 2="Reasonably Certain Underlying Cause of Death is Not Breast Cancer" 4="Questionable, But Underlying Cause of Death is Most Likely Not Breast Cancer" 5="Questionable, But Underlying Cause Of Death is Most Likely Breast Cancer"
mort_class_y	Mortality classification, reviewer y		.C="Not a Breast Cancer Case" 0="Not Reviewed" 1="Reasonably Certain Underlying Cause of Death is Breast Cancer" 2="Reasonably Certain Underlying Cause of Death is Not Breast Cancer" 4="Questionable, But Underlying Cause of Death is Most Likely Not Breast Cancer" 5="Questionable, But Underlying Cause Of Death is Most Likely Breast Cancer"
survival_stat	Survival status		.C="Not a Breast Cancer Case" 1="Living" 2="Deceased" 3="Lost To Follow-Up"

Section 9: Death

Variable	Label	Description	Format Text
adj_cod	Adjusted icd-8 cause of death		Numeric .N="Not Applicable"
adj_cod_cert	Adjusted cause of death from death certificate		Numeric .N="Not Applicable"
age_nonBCdeath	Age at death, if woman never had breast cancer		Numeric .N="Not Applicable"
anydeath_days	Days from randomization to date of death from any cause		Numeric .N="Not Applicable"
cod_deathcert	Underlying cause of death per death certificate		Numeric .N="Not Applicable"
cod_nonBC	4-digit icd-code from death certificate, if woman never had breast cancer		Numeric .N="Not Applicable"
cod_nonBC_icd8	4-digit icd code for cause of death, converted to icd8		Numeric .N="Not Applicable"
cod_reviewer	Underlying 3 digit cause of death per reviewers		Numeric .N="Not Applicable"
death_flag	Death flag		0="Alive" 1="Dead, Interval From Entry to Death <= 15 Years" 2="Dead, Interval From Entry to Death > 15 Years" 9="Status Unknown"
final_vital_stat	Vital status at end of follow-up (after death review)		0="Alive or Lost to Follow-Up" 1="Dead of Breast Cancer" 2="Dead of Cause Other Than Breast Cancer"
icdrev_nonBC	Icd revision at time of death, if woman never had breast cancer		Numeric .N="Not Applicable"
year_nonBCdeath_entry	Year of death following entry, if woman never had breast cancer		Numeric .N="Not Applicable"