GENIE BPC Analytic Data Guide NSCLC v2.0-public

Prepared by the Statistical Coordinating Center at Memorial Sloan Kettering Cancer Center

Contents

OVERVIEW

The GENIE Biopharma Collaborative (BPC) Core Team has created analytic datasets that include data from the AACR Project GENIE Tier 1 registry, institutions' tumor registries and electronic health records using the PRISSMM phenomic data model. Derived variables based on these data elements are also provided. This analytic data guide serves as a resource for all data elements included in the analytic datasets containing demographic and phenomic information, which can be combined with the genomic data that have been submitted to GENIE.

Release Notes: This document accompanies the **NSCLC v2.0-public** data release.

Abbreviations

AACR American Association for Cancer Research
BPC Biopharma Collaborative
CPT Cancer Panel Test
DFCI Dana-Farber Cancer Institute, Boston, MA, United States
EHR Electronic Health Record
GENIE Genomics Evidence Neoplasia Information Exchange
HIPAA The Health Insurance Portability and Accountability Act of 1996
MSK Memorial Sloan Kettering Cancer Center, New York, NY, United States
NAACCR North American Association of Central Cancer Registries
NDI National Death Index
NGS Next Generation Sequencing
NSCLC Non-Small Cell Lung Cancer
PRISSMM Pathology, **R**adiology, **I**maging, **S**igns, **S**ymptoms, Tumor **M**arkers, **M**edical Oncologist
QA Quality Assurance
UHN University Health Network Princess Margaret, Toronto, Ontario, Canada
VICC Vanderbilt-Ingram Cancer Center, Nashville, TN, United States

PRISSMM Overview and Note

PRISSMM is a system for extracting clinical data from longitudinal EHRs and for ascertaining high salience outcomes of cancer treatment in multi-site projects that span different health systems. PRISSMM includes a system for estimating "real world endpoints" from the text data contained in the EHR. PRISSMM is a flexible convenient consensus-based standard in the same way that RECIST criteria have become a standard for radiology endpoints in clinical trials. PRISSMM is the phenomic data that characterizes treatment exposures, and outcomes based on the results of pathology (P), imaging of local (R) and distant sites (I), signs (S), symptoms (S), tumor markers (M) and Medical Oncology assessments (M). The system specifies data provenance, timing for review, approaches to handling treatment gaps and discontinuity and endpoint specification that is applicable to longitudinal EHR data for patients with cancer.

PRISSMM training materials and this data guide are shared with the GENIE-BPC release. A license to the complete set of PRISSMM tools including databases used to extract data and training materials can be obtained by e-mailing PRISSMM@mskcc.org. The PRISSMM system was developed at DFCI and licensing is shared between MSK and DFCI. The materials are shared at no cost with academic medical centers and for a fee that is used to support development activities with for profit entities.

BPC Projects

GENIE BPC sponsors six projects in phase I that involve augmenting the genomic data in Project GENIE with PRISSMM phenomic data and cancer-related outcomes. Each project encompasses a single cancer site. The six projects are: bladder cancer, breast cancer, colon/rectal cancer, non-small cell lung cancer, pancreas cancer, and prostate cancer. A description of the GENIE BPC Project and its associated data quality assurance processes have been published (Lavery et al, 2022).

Eligibility

The BPC Project-specific eligibility criteria for the NSCLC project are as follows:

- Eligible OncoTree Diagnoses:
  - Non-Small Cell Lung Cancer (NSCLC)
  - Ciliated Muconodular Papillary Tumor of the Lung (CMPT)
  - Large Cell Lung Carcinoma (LCLC)
  - Lung Adenocarcinoma (LUAD)
  - Pleomorphic Carcinoma of the Lung (LUPC)
  - Lung Squamous Cell Carcinoma (LUSC)
  - Poorly Differentiated Non-Small Cell Lung Cancer (NSCLCPD)
  - Lung Adenosquamous Carcinoma (LUAS)
  - Mucoepidermoid Carcinoma of the Lung (LUMEC)
  - Lymphoepithelioma-like Carcinoma of the Lung (LECLC)
  - Clear Cell Carcinoma of the Lung (CCLC)
  - Large Cell Lung Carcinoma With Rhabdoid Phenotype (RLCLC)
  - Giant Cell Carcinoma of the Lung (GCLC)
  - Basaloid Large Cell Carcinoma of the Lung (BLCLC)
- Stage I-IV at diagnosis
- Genomic sequencing report at:
  - DFCI, MSK, or VICC between January 1, 2014 and December 31, 2017
  - UHN between January 1, 2014 and December 31, 2015
- Aged 18 or older at the time of genomic sequencing
- Minimum of two years of possible follow-up after sequencing

Cases meeting specified eligibility criteria are randomly selected from the AACR Project GENIE Cancer Registry.

Note that a patient can be selected for more than one project (e.g., if a patient is selected for the non-small cell lung cancer project, it is possible for that patient to also be selected for the breast cancer project if they met eligibility criteria for both projects).

These selection criteria may impact the generalizability of results since genomic sequencing is not always performed at diagnosis and therefore may lead to several forms of bias. Investigators who wish to test specific hypotheses should work with a statistician to perform analyses that account for these biases. Failure to do so may result in incorrect inferences (Brown et al, 2022).

Data Privacy

Compliance with data privacy required redaction of the name and duration of investigational drugs. Patients >89 years of age at any time point are excluded from this study cohort.

References

Lavery JA, Lepisto EM, Brown S, Rizvi H, McCarthy C, LeNoue-Newton M, Yu C, Lee J, Guo X, Yu T, Rudolph J, Sweeney S; AACR Project GENIE Consortium, Park BH, Warner JL, Bedard PL, Riely G, Schrag D, Panageas KS. A Scalable Quality Assurance Process for Curating Oncology Electronic Health Records: The Project GENIE Biopharma Collaborative Approach. JCO Clin Cancer Inform. 2022 Feb;6:e2100105. doi: 10.1200/CCI.21.00105. PMID: 35192403; PMCID: PMC8863125.

Brown S, Lavery JA, Shen R, Martin AS, Kehl KL, Sweeney SM, Lepisto EM, Rizvi H, McCarthy CG, Schultz N, Warner JL, Park BH, Bedard PL, Riely GJ, Schrag D, Panageas KS; AACR Project GENIE Consortium. Implications of Selection Bias Due to Delayed Study Entry in Clinical Genomic Studies. JAMA Oncol. 2022 Feb 1;8(2):287-291. doi: 10.1001/jamaoncol.2021.5153. PMID: 34734967.

Format

There are 8 analytic datasets included in the NSCLC v2.0-public data release. This data guide describes the variables included each dataset.

Dataset	File Name on Synapse
1. Patient Characteristics Dataset	patient_level_dataset.csv
2. BPC Project Cancer Diagnosis Dataset	cancer_level_dataset_index.csv
3. Non-BPC Project Cancer Diagnosis Dataset	cancer_level_dataset_non_index.csv
4. Cancer-Directed Regimen Dataset	regimen_cancer_level_dataset.csv
5. PRISSMM Pathology Dataset	pathology_report_level_dataset.csv
6. PRISSMM Imaging Dataset	imaging_level_dataset.csv
7. PRISSMM Medical Oncologist Assessment Dataset	med_onc_note_level_dataset.csv
8. Cancer Panel Test Dataset	cancer_panel_test_level_dataset.csv

We describe each variable included in these datasets using the following format:

Field name
[Variable name]
Value (character/numeric/date/date-time)
Description of the field
Data Standard (where applicable)

Variables have been color coded to help users understand their provenance.
Variables shown in orange indicate variables obtained from the AACR Project GENIE Tier 1 data.
Variables shown in green indicate variables obtained directly from the institution's tumor registry.
Variables shown in blue indicate variables obtained from curation of the EHR.
Variables shown in purple were derived by the Statistical Coordinating Center.

Field names shaded in gray indicate that another variable is preferred. For example, stage at diagnosis is captured from the tumor registry and also curated from the EHR, but a derived variable combining the two sources is also provided. In this case, the tumor registry and curated stage at diagnosis are shown with gray shading and the composite derived variable is preferred for analysis.

The following table provides further information about the types of variables.

Type of Variable	Description
AACR Project GENIE Tier 1 Data	Variables denoted in orange represent data that is obtained directly from the AACR Project GENIE Tier 1 Data. These variables correspond to the values stored for each sample in the Tier 1 Data. Users should refer to the AACR Project GENIE website for further detail.
Tumor registry at each institution	Variables denoted in green are obtained from each hospital's tumor registry system. These are NAACR-defined variables and thus are highly standardized across institutions. Not all patients have their data captured by the tumor registry. The provenance of the data is abstraction by Certified Tumor Registrars at each institution. The NAACCR data standards are noted for each registry-specific field and users should refer to the NAACCR website for further detail.
Curated	Variables denoted in blue have been curated according to the PRISSMM phenomic data curation model to extract unstructured data from the EHR.
Derived	Variables denoted in purple are calculated based on AACR Project GENIE Tier 1, tumor registry and/or curated variables. These variables are provided to facilitate ease of use as well as consistency across different analyses.

PATIENT CHARACTERISTICS DATASET

The Patient Characteristics dataset is structured as one record per patient.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, "GENIE"; the second component is the institution's abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Patient has any Redacted Time Interval Data to Comply with Health Insurance Portability and Accountability Act (HIPAA)

[redacted]

Value (Character)

- Yes
- No

Description

- Indicates whether any portion of the patient's data (across all datasets) was redacted to comply with HIPAA.
- Any time interval indicating age >89 years or that could be used to identify age >89 will not be available for patients who have [redacted] = "Yes"

Year of Birth

[birth_year]

Value (Numeric)

- YYYY

Description

- Patient's year of birth

Ethnicity: Spanish/Hispanic Origin

[naaccr_ethnicity_code]

Value (Character)

- Non-Spanish; non-Hispanic
- Mexican (includes Chicano)
- Puerto Rican
- Cuban
- South or Central American (except Brazil)
- Other specified Spanish/Hispanic origin (includes European; excludes Dominican Republic)
- Spanish NOS or Hispanic NOS or Latino NOS
- Spanish surname only
- Dominican Republic
- Unknown whether Spanish or not

Description

- Ethnicity of patient, independent of patient's race
- Institutions not collecting Spanish/Hispanic origin have set this field to "Unknown whether Spanish or not"

Data Standard: NAACCR #190

Race (Primary)

[naaccr_race_code_primary]

Value (Character)

- White
- Black
- American Indian, Aleutian, or Eskimo
- Chinese
- Japanese
- Filipino
- Hawaiian
- Korean
- Vietnamese
- Laotian
- Hmong
- Kampuchean (Cambodian)
- Thai
- Asian Indian or Pakistani NOS
- Asian Indian
- Pakistani
- Micronesian NOS
- Chamorro/Chamoru
- Guamanian NOS
- Polynesian NOS
- Tahitian
- Samoan
- Tongan
- Melanesian NOS
- Fiji Islander
- New Guinean
- Other Asian
- Pacific Islander NOS
- Other
- Unknown

Description

- First race specified, independent of ethnicity
- For institutions collecting more than one race category, this race variable indicates the primary race for the patient.
- Institutions not collecting race set this field to “Unknown”

Data Standard: NAACCR #160

Race (Secondary)

[naaccr_race_code_secondary]

Value (Character)

- White
- Black
- American Indian, Aleutian, or Eskimo
- Chinese
- Japanese
- Filipino
- Hawaiian
- Korean
- Vietnamese
- Laotian
- Hmong
- Kampuchean (Cambodian)
- Thai
- Asian Indian or Pakistani NOS
- Asian Indian
- Pakistani
- Micronesian NOS
- Chamorro/Chamoru
- Guamanian NOS
- Polynesian NOS
- Tahitian
- Samoan
- Tongan
- Melanesian NOS
- Fiji Islander
- New Guinean
- Other Asian
- Pacific Islander NOS
- No further race documented
- Other
- Unknown

Description

- Second race specified, independent of ethnicity
- Institutions not collecting secondary race set this field to “No further race documented”

Data Standard: NAACCR #161

Race (Tertiary)

[naaccr_race_code_tertiary]

Value (Character)

- White
- Black
- American Indian, Aleutian, or Eskimo
- Chinese
- Japanese
- Filipino
- Hawaiian
- Korean
- Vietnamese
- Laotian
- Hmong
- Kampuchean (Cambodian)
- Thai
- Asian Indian or Pakistani NOS
- Asian Indian
- Pakistani
- Micronesian NOS
- Chamorro/Chamoru
- Guamanian NOS
- Polynesian NOS
- Tahitian
- Samoan
- Tongan
- Melanesian NOS
- Fiji Islander
- New Guinean
- Other Asian
- Pacific Islander NOS
- No further race documented
- Other
- Unknown

Description

- Third race specified, independent of ethnicity
- Institutions not collecting tertiary race set this field to “No further race documented”

Data Standard: NAACCR #162

Race/Ethnicity

[race_ethnicity]

Value (Character)

- AAAPI (Asian, Asian American, and Pacific Islander)
- Asian Indian or Pakistani NOS
- Hispanic/Latinx
- Non-Hispanic Black
- Non-Hispanic White
- Other
- Unknown

Description

- Indicates the patient's race/ethnicity defined based on variables [naaccr_race_code_primary] and [naaccr_ethnicity_code]

Sex

[naaccr_sex_code]

Value (Character)

- Male
- Female
- Other intersex, disorders of sexual development/DSD
- Transsexual NOS
- Transsexual natal male
- Transsexual natal female

Description

- Patient's sex at time of diagnosis of index cancer

Data Standard: NAACCR #220

Time (Days) from Date of Birth to Date of Last Oncology Visit to Internal Institution

[last_oncvisit_int]

Value (Numeric)

Description

- Interval in days from date of birth to most recent date that the patient had an in-person or tele-visit visit with an oncology provider at the internal institution of care; these visits may include medical oncology, surgical oncology, radiation oncology, palliative care, social work, lab draws, imaging scans, emergency room, or hospital visits.
- Completeness of data is subject to availability of visit information at each participating institution. Consequently, the preferred variable for determining the date of last contact is [dob_lastalive_int].

Time (Days) from Date of Birth to Date Last Known Alive

[last_alive_int]

Value (Numeric)

Description

- Interval in days from date of birth to most recent date that the patient was known to be alive
- Requires documentation of a phone call or email exchange with the patient or a family member.
- Completeness of data is subject to availability of visit information at each participating institution. Consequently, the preferred variable for determining the date of last contact is [dob_lastalive_int].

Time (Days) from Date of Birth to Date of Last Known Non-Oncology Visit to Internal Institution

[last_anyvisit_int]

Value (Numeric)

Description

- Interval in days from date of birth to most recent date that the patient had an in-person or tele-visit visit to the internal institution for non-oncology related care, including visits with primary care or cardiology.
- Completeness of data is subject to availability of visit information at each participating institution. Consequently, the preferred variable for determining the date of last contact is [dob_lastalive_int].

Time (Days) from Date of Birth to Enrollment in Hospice Care

[enroll_hospice_int]

Value (Numeric)

Description

- Interval in days from date of birth to the date that the patient was enrolled in hospice care.
- Completeness of data is subject to availability of visit information at each participating institution. Consequently, the preferred variable for determining the date of last contact is [dob_lastalive_int].

Time (Days, Months, Years) from Date of Birth to Last Known Alive Date

[dob_lastalive_int], [dob_lastalive_int_mos], [dob_lastalive_int_yrs]

Value (Numeric)

Description

- Interval in days [dob_lastalive_int], months [dob_lastalive_int_mos], and years [dob_lastalive_int_yrs] from date of birth to last known alive date.
- This variable is recommended as the censoring date in survival analyses.
- Based on the composite of the most recent date that the patient: 1. received oncology care at the internal institution, 2. received any care at the internal institution, 3. documentation from any source that the patient is alive, and 4. the patient is enrolled in hospice.

Age (Years) at Last Known Alive Date

[age_last_fu_yrs]

Value (Numeric)

Description

- Age of the patient at the time last known to be alive.
- Based on the composite of the most recent date that the patient: 1. received oncology care at the internal institution, 2. received any care at the internal institution, 3. documentation from any source that the patient is alive, and 4. the patient is enrolled in hospice.

Time (Days) from Date of Birth to Death

[hybrid_death_int]

Value (Numeric)

Description

- Interval in days from date of birth to date of death
- Populated only if patient is known to be dead at time of curation

Age (Years) at Death Date

[age_death_yrs]

Value (Numeric)

Description

- Age at death, in years
- Populated only if patient is known to be dead at the time of curation

Source of Death Information

[hybrid_death_source]

Value (Character)

- Curated
- EHR
- NDI
- Tumor Registry
- Other

Description

- Indicates source of death information
- Populated only if patient is known to be dead at the time of curation

Number of Cancers, Any Type

[n_cancers]

Value (Numeric)

Description

- Number of invasive and non-invasive/in situ cancer diagnoses ever experienced by the patient
- Based on the count of records in Cancer Diagnosis dataset for each record ID
- Basal cell and squamous cell skin cancers are not included

Number of BPC Project Cancers (Index Cancers)

[n_cancers_index]

Value (Numeric)

Description

- Number of BPC Project Cancers that were identified for a patient
- A BPC Project Cancer is defined as the cancer that met eligibility criteria, underwent genomic sequencing and was submitted to AACR Project GENIE.
- Each patient has at least one BPC Project Cancer. Patients may have multiple BPC Project Cancers.
- PRISSMM data elements are curated for BPC Project Cancers.
- The terms "BPC Project Cancer" and "index cancer" are used interchangeably.
- Further details regarding the definition of BPC Project and non-BPC Project Cancers can be found in Appendix 1: BPC Project and Non-BPC Project Cancers

Number of Cancer-Directed Regimens Curated

[n_regimens_pt]

Value (Numeric)

Description

- The total number of cancer-directed regimens, including anti-neoplastic, immunotherapy, and hormone therapy that the patient has ever received for any cancer diagnosis
- Based on the count of records in the Cancer-Directed Regimen dataset for each record ID
- This number includes cancer-directed regimens given for non-BPC Project Cancers.

Number of Imaging Reports

[n_imaging_reports_pt]

Value (Numeric)

Description

- The total number of imaging scans curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of records in the Imaging dataset for each record ID
- Imaging scans include: CTs, MRIs, PET, PET/CTs, Bone Scans, and Nuclear Medicine scans
- This number includes scans performed for non-BPC Project Cancers
- For patients without a diagnosis of cancer in the breast, mammograms were not reviewed. For patients with a diagnosis of cancer in the breast, mammograms are reviewed starting at the month/year of the BPC Project Cancer diagnosis.

Number of CT Scans

[n_scans_ct_pt]

Value (Numeric)

Description

- The number of CT imaging scans curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of CT scans in the Imaging dataset for each record ID
- This number includes CT scans performed for non-BPC Project Cancers

Number of MRIs

[n_scans_mri_pt]

Value (Numeric)

Description

- The number of MRI imaging scans curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of MRIs in the Imaging dataset for each record ID
- This number includes MRIs performed for non-BPC Project Cancers

Number of PET or PET-CT Scans

[n_scans_pet_ct_pt]

Value (Numeric)

Description

- The number of PET or PET-CT imaging scans curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of PET and PET-CT scans in the Imaging dataset for each record ID
- This number includes PET or PET-CT scans performed for non-BPC Project Cancers

Number of Mammograms (Breast Cancer Only)

[n_scans_mammog_pt]

Value (Numeric)

Description

- The number of mammograms curated starting in the month/year of the BPC Project Cancer diagnosis for patients with any (BPC Project or non-BPC Project) diagnosis of breast cancer
- Based on the count of mammograms in the Imaging dataset for each record ID
- This number includes mammograms performed for non-BPC Project Cancers of the breast.
- For patients without a diagnosis of cancer in the breast, mammograms were not reviewed

Number of Bone Scans

[n_scans_bone_pt]

Value (Numeric)

Description

- The number of bone scans curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of bone scans in the Imaging dataset for each record ID
- This number includes bone scans performed for non-BPC Project Cancers

Number of Other Scans

[n_scans_other_pt]

Value (Numeric)

Description

- The number of other imaging scans, including other nuclear medicine scans, curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of other scans in the Imaging dataset for each record ID
- This number includes scans performed for non-BPC Project Cancers

Number of Pathology Reports

[n_path_reports_pt]

Value (Numeric)

Description

- The number of pathology reports starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of records in the Pathology dataset for each record ID
- This number includes pathology reports for non-BPC Project Cancers

Number of Medical Oncologist Assessments Curated

[n_md_notes_pt]

Value (Numeric)

Description

- The number of medical oncologist assessments curated starting in the month/year of the BPC Project Cancer diagnosis
- Based on the count of records in the Medical Oncologist Assessment dataset for each record ID
- One medical oncologist assessment per month was curated. Curation instructions are provided in Appendix 3.

Number of Eligible Cancer Panel Tests Curated

[n_cpt_pt]

Value (Numeric)

Description

- The number of cancer panel tests that met the eligibility criteria and were curated
- Based on the count of records in the Cancer Panel Test dataset
- Not all sequenced specimens included in the AACR Project GENIE repository appear in this dataset due to eligibility requirements.

Release Version Number

[release_version]

Value (Character)

- 2.0-public

Description

- Indicates the version number of the data release

CANCER-DIAGNOSIS DATASETS

Two Cancer Diagnosis datasets are provided.

1. The BPC Project Cancer Diagnosis dataset contains one record per BPC Project Cancer diagnosis, per patient. A BPC Project Cancer is the cancer that met the eligibility criteria for the project and was selected for PRISSMM phenomic data curation.

This dataset can be linked to the following datasets:

- *Cancer-Directed Regimen dataset and Cancer Panel Test dataset using variables [cohort], [record_id] and [ca_seq]*
- *Patient Characteristics, PRISSMM Pathology, PRISSMM Imaging, and PRISSMM Medical Oncologist Assessment datasets using [cohort] and [record_id].*

2. The Non-BPC Project Cancer Diagnosis dataset contains one record per non-BPC Project Cancer diagnosis, per patient. This dataset includes two types of cancer diagnoses: 1) non-BPC Project invasive cancer and in situ/non-invasive cancer diagnoses, and 2) other tumors.

This dataset can be linked to the following datasets:

- *Cancer-Directed Regimen dataset using variables [record_id] and [ca_seq]*
- *Patient Characteristics, PRISSMM Pathology, PRISSMM Imaging, and PRISSMM Medical Oncologist Assessment datasets using [record_id].*
- *Cannot be linked to the cancer panel test dataset because non-BPC Project Cancer diagnoses were not genomically sequenced (Appendix 1)*

Further details regarding the definition of BPC Project and non-BPC Project Cancers can be found in Appendix 1.

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The recommended variable is noted in the description.

Note that NAACCR variables are not available for patients from UHN.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, "GENIE"; the second component is the institution's abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Diagnosis Eligible for Curation

[tr_eligible]

Value (Numeric)

- 1 = Eligible for full curation
- 0 = Ineligible for full curation

Description

- Indicates whether the diagnosis is eligible for curation from tumor registry. A cancer diagnosis is eligible for curation if the behavior code [naaccr_behavior_cd] is 2 (non-invasive/in situ) or 3 (invasive/malignant). Cancer diagnoses with behavior codes 0 or 1 are not eligible for full annotation in the BPC project, but the tumor registry data are available.

Cancer Sequence Identifier

[ca_seq]

Value (Numeric)

- 0 = first and only cancer
- 1 = first of two or more primaries
- 2 = second of two or more primaries
- …10 = tenth of ten or more primaries

Description

- Sequential order of cancer diagnoses based on the date of diagnosis
- Primary key for the Cancer Diagnosis, Cancer-Directed Regimen and Cancer Panel Test datasets

BPC Project (Index) Cancer Indicator

[redcap_ca_index]

Value (Character)

- Yes
- No

Description

- The BPC Project Cancer is defined as the cancer that met eligibility criteria, underwent genomic sequencing and was submitted to AACR Project GENIE.
- Each patient has at least one BPC Project Cancer. Patients may have multiple BPC Project Cancers, though this is rare.
- PRISSMM data elements are curated for BPC Project Cancers.
- The terms “BPC Project Cancer” and “index cancer” are used interchangeably.
- Further details regarding the definition of BPC Project and non-BPC Project Cancers can be found in Appendix 1. BPC Project and Non-BPC Project Cancers.

Time (Days, Months, Years) from Date of Birth to Cancer Diagnosis

[dob_ca_dx_days], [dob_ca_dx_mos], [dob_ca_dx_yrs]

Value (Numeric)

Description

- Interval in days [dob_ca_dx_days], months [dob_ca_dx_mos], or years [dob_ca_dx_yrs] between date of birth and date of cancer diagnosis

Source of Cancer Diagnosis Date

[ca_dx_how]

Value (Character)

- Pathology
- Imaging
- Physical Exam
- Other

Description

- Source of the curated diagnosis date
- If [ca_dx_how] = "Pathology", the cancer diagnosis date is based upon a review of the pathology report indicating the first histologic confirmation of cancer.
- Populated only if diagnosis is eligible for curation ([tr_eligible] = 1)

Time (Days, Months, Years) from Date of Birth to Next Curated Cancer Diagnosis

[dob_next_ca_days], [dob_next_ca_mos], [dob_next_ca_yrs]

Value (Numeric)

Description

- Time in days [dob_next_ca_days], months [dob_next_ca_mos], or years [dob_next_ca_yrs] from date of birth to the next curated cancer diagnosis
- Populated only if patient has a subsequent cancer diagnosis

Time (Days, Months, Years) to First BPC Project Cancer Diagnosis

[first_index_ca_days], [first_index_ca_mos], [first_index_ca_yrs]

Value (Numeric)

Description

- For patients with multiple cancer diagnoses, the time in days [first_index_ca_days], months [first_index_ca_mos], or years [first_index_ca_yrs] to first BPC Project Cancer diagnosis from other cancer diagnosis
- Time is negative if the BPC Project Cancer occurred prior to the comparative cancer; time is positive if BPC Project Cancer occurred after the comparative cancer
- Populated only if cancer diagnosis is not the first BPC Project Cancer. This field is blank for the first BPC Project Cancer

Tumor Registry Time (Days) from Date of Birth to First Contact at Institution

[naaccr_first_contact_int]

Value (Numeric)

Description

- Interval in days from date of birth to first contact at institution based on tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #580

Age (Years) at Diagnosis

[age_dx]

Value (Numeric)

Description

- Patient age at diagnosis

Primary Cancer Site

[ca_d_site]

Value (Character)

- ICD-O-3 topography code
- F10 = Cerebral Spinal Fluid
- F20 = Peritoneal Fluid/Ascites
- F30 = Pleural Fluid
- F40 = Urine
- F50 = Pericardial Fluid

Description

- ICD-O-3 code for the primary cancer site

Data Standard: ICD-O-3 Topography Codes

Cancer Type

[ca_type]

Value (Character)

- Adrenocortical Carcinoma
- Anal Cancer
- Appendix Cancer
- Bile Duct Cancer
- Bladder Cancer
- Brain Cancer
- Breast Cancer
- NET or Carcinoid
- Cervical Cancer
- Colon Cancer
- Colon/Rectum Cancer
- Esophagus Cancer
- Ewing Sarcoma
- Fallopian Tube Cancer
- Gallbladder Cancer
- Germ Cell Tumor
- GIST
- Head and Neck Cancer
- Mesothelioma
- Ill Defined/Cancer of Unknown Primary
- Liver Cancer
- Lung Cancer, NOS
- Melanoma
- Merkel Cell
- Neuroblastoma
- Non Small Cell Lung Cancer
- Osteosarcoma
- Ovarian Cancer
- Pancreatic Cancer
- Parathyroid Cancer
- Penis Cancer
- Peritoneum Cancer
- Placenta Cancer
- Prostate Cancer
- Rectum and Rectosigmoid Cancer
- Renal Kidney Cancer
- Renal Pelvis Cancer
- Retinoblastoma
- Rhabdomyosarcoma
- Scrotum Cancer
- Small Cell Lung Cancer
- Small Intestine Cancer
- Stomach Cancer
- Testis Cancer
- Thymus Cancer
- Thyroid Cancer
- Uterus Cancer
- Vagina Cancer
- Vulva Cancer
- Wilms Tumor
- Other

Description

- Cancer type was characterized based on information that includes ICD-O-3 topography and morphology codes.
- Mappable to AJCC Collaborative Stage v2.05, AJCC staging v7 and v8, SEER, and NCI.
- Populated only if diagnosis is not a hematopoietic or lymphoid neoplasm or pre-malignancy ([ca_heme_malig] = “No”)
- Note: this variable is largely missing for this cohort

Brain Cancer Type

[ca_type_brain]

Value (Character)

- Astrocytic Tumors
- Diffuse astrocytic and oligodendroglial tumors
- Choriod plexus tumors
- Craniopharyngioma (Grade I)
- Embryonal tumors
- Ependymal Tumors
- Germ cell tumors
- Medulloblastomas
- Meningeal Tumors
- Mesenchymal non meningothelial tumors
- Mixed gliomas
- Neuronal and mixed neuronal-gilial tumors
- Oligodendroglial tumors
- Other astrocytic tumors
- Other gliomas
- Pineal Parenchymal Tumors
- Tumors of cranial and paraspinal nerves
- Tumors of the pineal region
- Tumors of the sellar region
- Other

Description

- Populated only if:
  - Cancer type is recorded as brain cancer ([ca_type] = “Brain Cancer”)
  - Non-BPC Project Cancer

Data Standard: World Health Organization Central Nervous System Tumor Classification System (2018)

Cancer Type, Other

[ca_type_oth]

Value (Character)

- Free-text

Description

- Populated only if:
  - Cancer type is recorded as “other” ([ca_type] = “Other”)
  - Non-BPC Project Cancer

Hematopoietic or Lymphoid Neoplasm or Pre-Malignancy

[ca_heme_malig]

Value (Character)

- Yes
- No

Description

- Indicates whether cancer diagnosis is a hematologic malignancy
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Non-BPC Project Cancer

Type of Hematopoietic or Lymphoid Neoplasm

[ca_heme_type]

Value (Character)

- Leukemia
- Lymphoma
- Langerhans Cell Histiocytosis
- MDS Myelodysplastic Syndrome
- MGUS Monoclonal gammopathy of undetermined significance
- Multiple Myeloma
- Plasmacytoma
- Other hematopoietic or lymphoid neoplasm
- Unspecified

Description

- Populated only if:
    - Hematopoietic or lymphoid neoplasm or pre-malignancy ([ca_heme_malig] = “Yes”).
    - Non-BPC Project Cancer

Data Standard: National Cancer Institute Cancer Types: Hematologic/Blood

Tumor Registry Histology

[naaccr_histology_cd]

Value (Numeric)

- ICD-O-3 morphology code

Description

- Histology code from tumor registry (ICD-O-3)
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #552

Histology Category

[ca_hist_adeno_squamous]

Value (Character)

- Adenocarcinoma
- Carcinoma
- Squamous cell
- Sarcoma
- Small cell carcinoma
- Other histologies/mixed tumor

Description

- Broad histology group based on [naaccr_histology_cd], if available. If unavailable, may be derived using information in the pathology report dataset.
- Populated only if BPC Project Cancer

Tumor Registry ICD-O-3 Behavior Code

[naaccr_behavior_cd]

Value (Numeric)

- 0 = Benign
- 1 = Borderline
- 2 = In situ and/or carcinoma in situ
- 3 = Invasive

Description

- Behavior code from tumor registry (ICD-O-3)
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #523

Tumor Registry Laterality Code

[naaccr_laterality_cd]

Value (Character)

- 0 = Not a paired site
- 1 = Right: origin of primary
- 2 = Left: origin of primary
- 3 = Only one side involved, right or left origin unspecified
- 4 = Bilateral involvement at time of diagnosis, lateral origin unknown for a single primary; or both ovaries involved simultaneously, single histology; bilateral retinoblastomas; bilateral Wilms' tumors
- 5 = Paired site: midline tumor
- 9 = Paired site, but no information concerning laterality
- Not paired

Description

- Laterality code from the tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #410

Grade or Differentiation of Tumor

[ca_grade]

Value (Character)

- I
- II
- III
- IV
- B-cell
- T-cell

Description

- The grade or degree of differentiation of the tumor

Data Standard: NAACCR #440

Tumor Registry Best Group Stage

[best_ajcc_stage_cd]

Value (Character)

- Free-text

Description

- Best stage group calculated by each institution's tumor registry software (METRIQ Tumor Registry Algorithm)
- Populated only if:
    - Cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.
    - BPC Project Cancer diagnosis
- Preferred derived variable [stage_dx] incorporates both tumor registry and curated stage.

Curated Stage IV at Diagnosis

[ca_stage_iv]

Value (Character)

- No
- Yes
- Not Applicable
- Unknown

Description

- Indicates whether cancer was diagnosed as stage IV
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry, unless the tumor registry Best Group Stage is not available [best_ajcc_stage_cd] ≠ 88, 99)
    - Diagnosis is not a hematopoietic or lymphoid neoplasm or pre-malignancy ([ca_heme_malig] = "No").
- Preferred derived variable [stage_dx] incorporates both tumor registry and curated stage.

Curated Group Stage at Diagnosis

[ca_stage]

Value (Character)

- 0
- 0A
- 0is
- I
- IA
- IB
- IC
- II
- IIA
- IIB
- IIC
- III
- IIIA
- IIIB
- IIIC
- IV
- IVA
- IVB
- IVC
- Not Applicable
- Unknown

Description

- Curated stage group at diagnosis
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Cancer diagnosis was not diagnosed at stage IV ([ca_stage_iv] = “No”).
- Preferred derived variable [stage_dx] incorporates both tumor registry and curated stage.

Derived Stage at Diagnosis

[stage_dx]

Value (Character)

- Stage 0
- Stage I
- Stage II
- Stage III
- Stage I-III NOS
- Stage IV

Description

- Cancer stage at diagnosis
- Derived from a combination of tumor registry and curated stage variables based on the following hierarchy:
    - Tumor registry best group stage [best_ajcc_stage_cd] based on the AJCC version documented [naaccr_tnm_edition_num]
    - If tumor registry best group stage was unavailable, curated group stage [ca_stage_iv], [ca_stage] based on AJCC version 7 was used
    - If tumor registry and curated group stage were both unavailable:
        - If patient received neoadjuvant chemotherapy or radiation therapy before pathological stage diagnosis, then clinical staging was used, if available. If unavailable, then pathologic staging was used
        - If patient did not receive neoadjuvant chemotherapy or radiation therapy before pathological stage diagnosis, then pathologic staging was used, if available. If unavailable, then clinical staging was used
- Populated for BPC Project Cancers only

Derived Stage IV at Diagnosis

[stage_dx_iv]

Value (Character)

- Stage 0
- Stage I-III
- Stage IV

Description

- Grouped cancer stage at diagnosis
- Derived from tumor registry best group stage [best_ajcc_stage_cd], if available. If unavailable, curated group stage was used [ca_stage_iv].

Tumor Registry TNM Pathologic Stage

[naaccr_tnm_path_desc]

Value (Character)

- Free-text

Description

- TNM pathology from tumor registry
  - T describes the size of the tumor and any spread of cancer into nearby tissue
  - N describes spread of cancer to nearby lymph nodes
  - M describes metastasis
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #920

Tumor Registry Pathologic T Stage

[naaccr_path_t_cd]

Value (Character)

- Free-text

Description

- Pathologic T stage from tumor registry
- T describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #880

Curated Pathologic T Stage

[ca_path_t_stage]

Value (Character)

- TX
- T0
- T1
- T2
- T3
- T4
- Not Applicable
- Unknown

Description

- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
  - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
  - Cancer was not diagnosed at stage IV ([ca_stage_iv] = “No”).

Data Standard: NAACCR #880

Curated Pathologic T1 Stage Detail

[ca_path_t1_det]

Value (Character)

- T1mic
- T1a
- T1a2
- T1b
- T1b1
- T1b2
- T1c
- T1d
- Not Applicable
- Unknown

Description

- Curated pathologic T1 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic T stage is T1 ([ca_path_t_stage] = “T1”).

Data Standard: NAACCR #880

Curated Pathologic T2 Stage Detail

[ca_path_t2_det]

Value (Character)

- T2a
- T2a1
- T2a2
- T2b
- T2c
- T2d
- Not Applicable
- Unknown

Description

- Curated pathologic T2 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic T stage is T2 ([ca_path_t_stage] = “T2”).

Data Standard: NAACCR #880

Curated Pathologic T3 Stage Detail

[ca_path_t3_det]

Value (Character)

- T3a
- T3b
- T3c
- T3d
- Not Applicable
- Unknown

Description

- Curated pathologic T3 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic T stage is T3 ([ca_path_t_stage] = “T3”).

Data Standard: NAACCR #880

Curated Pathologic T4 Stage Detail

[ca_path_t4_det]

Value (Character)

- T4a
- T4b
- T4c
- T4d
- T4e
- Not Applicable
- Unknown

Description

- Curated pathologic T4 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic T stage is T4 ([ca_path_t_stage] = "T4").

Data Standard: NAACCR #880

Tumor Registry Pathologic N Stage

[naaccr_path_n_cd]

Value (Character)

- Free-text

Description

- Pathologic N stage from tumor registry
- N describes spread of cancer to nearby lymph nodes
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #890

Curated Pathologic N Stage

[ca_path_n_stage]

Value (Character)

- NX
- N0
- N1
- N2
- N3
- N4
- Not Applicable
- Unknown

Description

- Curated pathologic N stage
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Cancers was not diagnosed at stage IV ([ca_stage_iv] = “No”).

Data Standard: NAACCR #890

Curated Pathologic N1 Stage Detail

[ca_path_n1_det]

Value (Character)

- N1mi
- N1a
- N1b
- N1c
- Not Applicable
- Unknown

Description

- Curated pathologic N1 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic N stage is N1 ([ca_path_n_stage] = “N1”).

Data Standard: NAACCR #890

Curated Pathologic N2 Stage Detail

[ca_path_n2_det]

Value (Character)

- N2a
- N2b
- N2c
- Not Applicable
- Unknown

Description

- Curated pathologic N2 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic N stage is N2 ([ca_path_n_stage] = “N2”).

Data Standard: NAACCR #890

Curated Pathologic N3 Stage Detail

[ca_path_n3_det]

Value (Character)

- N3a
- N3b
- N3c
- Not Applicable
- Unknown

Description

- Curated pathologic N3 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Pathologic N stage is N3 ([ca_path_n_stage] = “N3”).

Data Standard: NAACCR #890

Tumor Registry Pathologic M Stage

[naaccr_path_m_cd]

Value (Character)

- Free-text

Description

- Pathologic M stage from tumor registry
- M describes metastasis
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #900

Tumor Registry Pathologic Group Stage

[naaccr_path_stage_cd]

Value (Character)

- Free-text

Description

- Pathologic group stage from tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #910

Curated Pathologic Group Stage

[ca_path_group_stage]

Value (Character)

- 0
- 0A
- 0is
- I
- IA
- IA1
- IA2
- IB
- IB1
- IB2
- IC
- IS
- II
- IIA
- IIA1
- IIA2
- IIB
- IIC
- III
- IIIA
- IIIB
- IIIC
- IIIC1
- IIIC2
- IV
- IVA
- IVA1
- IVA2
- IVB
- IVC
- Occult
- Not Applicable
- Unknown

Description

- Populated only if:
  - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
  - Cancer was not diagnosed at stage IV ([ca_stage_iv] = “No”).

Data Standard: NAACCR #910

Tumor Registry Clinical T Stage

[naaccr_clin_t_cd]

Value (Character)

- Free-text

Description

- Clinical T stage from tumor registry
- T describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #940

Curated Clinical T Stage

[ca_clin_t_stage]

Value (Character)

- TX
- T0
- T1
- T2
- T3
- T4
- Not Applicable
- Unknown

Description

- Curated clinical T stage
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - If cancer was not diagnosed at stage IV ([ca_stage_iv] = “No”).

Data Standard: NAACCR #940

Curated Clinical T1 Stage Detail

[ca_clin_t1_det]

Value (Character)

- T1mic
- T1a
- T1a2
- T1b
- T1b1
- T1b2
- T1c
- T1d
- Not Applicable
- Unknown

Description

- Curated clinical T1 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical T stage is T1 ([ca_clin_t_stage] = “T1”).

Data Standard: NAACCR #940

Curated Clinical T2 Stage Detail

[ca_clin_t2_det]

Value (Character)

- T2a
- T2a1
- T2a2
- T2b
- T2c
- T2d
- Not Applicable
- Unknown

Description

- Curated clinical T2 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical T stage is T2 ([ca_clin_t_stage] = “T2”).

Data Standard: NAACCR #940

Curated Clinical T3 Stage Detail

[ca_clin_t3_det]

Value (Character)

- T3a
- T3b
- T3c
- T3d
- T3e
- Not Applicable
- Unknown

Description

- Curated clinical T3 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical T stage is T3 ([ca_clin_t_stage] = “T3”).

Data Standard: NAACCR #940

Curated Clinical T4 Stage Detail

[ca_clin_t4_det]

Value (Character)

- T4a
- T4b
- T4c
- T4d
- T4e
- Not Applicable
- Unknown

Description

- Curated clinical T4 stage detail
- T stage describes the size of the tumor and any spread of cancer into nearby tissue
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical T stage is T4 ([ca_clin_t_stage] = “T4”).

Data Standard: NAACCR #940

Tumor Registry Clinical N Stage

[naaccr_clin_n_cd]

Value (Character)

- Free-text

Description

- Clinical N stage from tumor registry
- N describes spread of cancer to nearby lymph nodes
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #950

Curated Clinical N Stage

[ca_clin_n_stage]

Value (Character)

- NX
- N0
- N1
- N2
- N3
- N4
- Not Applicable
- Unknown

Description

- Curated clinical N stage
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Cancer was not diagnosed at stage IV ([ca_stage_iv] = "No").

Data Standard: NAACCR #950

Curated Clinical N1 Stage Detail

[ca_clin_n1_det]

Value (Character)

- N1a
- N1b
- N1c
- Not Applicable
- Unknown

Description

- Curated clinical N1 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical N stage is N1 ([ca_clin_n_stage] = "N1").

Data Standard: NAACCR #950

Curated Clinical N2 Stage Detail

[ca_clin_n2_det]

Value (Character)

- N2a
- N2b
- N2c
- Not Applicable
- Unknown

Description

- Curated clinical N2 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical N stage is N2 ([ca_clin_n_stage] = “N2”).

Data Standard: NAACCR #950

Curated Clinical N3 Stage Detail

[ca_clinical_n3_det]

Value (Character)

- N3a
- N3b
- N3c
- Not Applicable
- Unknown

Description

- Curated clinical N3 stage detail
- N describes spread of cancer to nearby lymph nodes
- Populated only if:
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Curated clinical N stage is N3 ([ca_clin_n_stage] = “N3”).

Data Standard: NAACCR #950

Tumor Registry Clinical M Stage

[naaccr_clin_m_cd]

Value (Character)

- Free-text

Description

- Clinical M stage from tumor registry
- M describes metastasis
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #960

Tumor Registry Clinical Group Stage

[naaccr_clin_stage_cd]

Value (Character)

- Free-text

Description

- Clinical group stage from tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #970

Tumor Registry General Summary Stage

[naaccr_seer_sum_stage]

Value (Character)

- Free-text

Description

- General summary stage from tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #759

Neoadjuvant Chemotherapy or Radiation Therapy Before Pathologic Stage Diagnosis

[ca_tx_pre_path_stage]

Value (Character)

- Yes
- No
- Not Applicable
- Unknown

Description

- Indicates whether patient received neoadjuvant chemotherapy or radiation therapy before pathologic stage diagnosis
- Populated only if cancer was not diagnosed at stage IV ([ca_stage_iv] = "No")

Data Standard: NAACCR #920

Curated General Summary Stage

[ca_gen_sum_stage_2]

Value (Character)

- Unstaged
- In situ
- Localized
- Regional direct extension only
- Regional lymph nodes only
- Regional direct extension and regional lymph nodes
- Regional NOS
- Distant
- Not applicable

Description

- Summary stage includes all information through completion of surgery in the first course of treatment or within 4 months of diagnosis in the absence of disease progression, whichever is longer.
- Populated only if:
    - Cancer is non-BPC Project Cancer
    - Cancer diagnosis is curated. This field is blank for cancer diagnoses obtained from the tumor registry.
    - Cancer type is Neuroblastoma, Brain Cancer, Ewing Sarcoma, Retinoblastoma, Rhabdomyosarcoma, or Wilms Tumor ([ca_type] = "Neuroblastoma", "Brain Cancer", "Ewing Sarcoma", "Retinoblastoma", "Rhabdomyosarcoma", or "Wilms Tumor").

Data Standard: NAACCR #759

Tumor Registry TNM Edition Number

[naaccr_tnm_edition_num]

Value (Character)

- Free-text

Description

- TNM edition number corresponding to stage variables from tumor registry
- Populated only if cancer diagnosis is obtained from the tumor registry. This field is blank for curated cancer diagnoses.

Data Standard: NAACCR #1060

Presence of Distant Metastasis at the Time of Cancer Diagnosis (Stage IV Diagnoses)

[ca_dmets_yn]

Value (Character)

- Yes
- No - patient is stage IV with no distant metastases
- Unknown or Not mentioned

Description

- Indicates whether stage IV patient had distant metastases at time of cancer diagnosis
- Populated only if cancer diagnosed at stage IV ([best_ajcc_stage_cd] = "4", "4A" "4B" or [ca_stage_iv] = "Yes")

Sites of Distant Metastases at Cancer Diagnosis (Stage IV Diagnoses)

[ca_first_dmets1]-[ca_first_dmets10]

Value (Character)

- ICD-O-3 topography code

Description

- Site of distant metastases at diagnosis
- Up to 10 sites of distant metastasis at diagnosis are recorded
- Populated only if distant metastases are present at time of diagnosis ([ca_dmets_yn] = "Yes")

Non-Small Cell Lung Cancer: Cigarette Use at Time of Diagnosis

[ca_lung_cigarette]

Value (Character)

- Never used
- Current user
- Former user (quit <1 year)
- Former user (quit >1 year)
- Former user (unknown time)
- Unknown

Description

- Smoking status of patient at time of NSCLC diagnosis
- Populated only if NSCLC diagnosis (BPC Project and non-BPC Project Cancer diagnoses).

Data Standard: National Program of Cancer Registries (Reference: Capture of tobacco use among population-based registries: Findings from 10 National Program of Cancer Registries states)

Non-Small Cell Lung Cancer: Separate Tumor Nodules

[ca_lung_sep_tumor]

Value (Character)

- No separate tumor nodules
- Separate tumor nodules of same histologic type in ipsilateral lung same lobe
- Separate tumor nodules of same histologic type in ipsilateral lung, different lobe
- Separate tumor nodules of same histologic type in ipsilateral lung same AND different lobes
- Separate tumor nodules of same histologic type in ipsilateral lung unknown if same or different lobe(s)
- Multiple nodules or foci of tumor present, not classifiable based on directives in guide
- Not applicable
- Unknown if separate tumor nodules

Description

- Separate tumor nodules are defined as intrapulmonary metastasis identified in the same lobe or same lung (ipsilateral) originating from a single lung primary at the time of diagnosis. Biopsy of tumors may or may not be performed.
- Populated only if NSCLC diagnosis (BPC Project and non-BPC Project Cancer diagnoses).

Data Standard: NAACCR #2880

Non-Small Cell Lung Cancer: Pleural/Elastic Layer Invasion (PL)

[ca_lung_pl_el_inv]

Value (Character)

- PL 0; No evidence of visceral pleural invasion (PL)
- PL 1; Invasion beyond the visceral elastic pleura, but limited to the pulmonary pleura
- PL 2; Invasion to the surface of the pulmonary pleura
- PL 3; Tumor extends to the parietal pleura
- Invasion of pleura
- No histologic examination of pleura to assess pleural layer invasion
- Unknown if PL present; PL/elastic layer cannot be assessed; Not documented in patient record

Description

- Pleural/Elastic Layer Invasion (PL) by Hematoxylin and Eosin stain (H & E) or Electric Strain
- Populated only if NSCLC diagnosis (BPC Project and non-BPC Project Cancer diagnoses).

Data Standard: NAACCR #2890

Number of Regimens Associated with the Cancer Diagnosis

[ca_n_regimens]

Value (Numeric)

Description

- Count of cancer-directed regimens that were associated with the cancer diagnosis

Distant Metastasis Post Diagnosis (Stage I-III)

[dmets_stage_i_iii]

Value (Numeric)

- 0 = No distant metastasis
- 1 = Distant metastasis

Description

- Indicates the presence of distant metastasis post-diagnosis according to a radiology or pathology report
- Distant metastasis classified based on Appendix Table 5
- Populated only among patients stage I-III at diagnosis and if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis of Stage I-III to Distant Metastasis

[dx_to_dmets_days], [dx_to_dmets_mos], [dx_to_dmets_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_days]; months [dx_to_dmets_mos]; or years [dx_to_dmets_yrs] from diagnosis of stage I-III BPC Project Cancer to date of first indication of distant metastasis
- Populated only among patients stage I-III at diagnosis with distant metastasis post diagnosis ([dmets_stage_i_iii] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Abdomen

[dmets_abdomen]

Value (Numeric)

- 0 = No distant metastasis in the abdomen
- 1 = Distant metastasis in the abdomen

Description

- Indicates the presence of distant metastasis in the abdomen post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the abdomen post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the abdomen at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Abdomen

[dx_to_dmets_abdomen_days], [dx_to_dmets_abdomen_mos], [dx_to_dmets_abdomen_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_abdomen_days]; months [dx_to_dmets_abdomen_mos]; or years [dx_to_dmets_abdomen_yrs] from diagnosis to date of first distant metastasis in the abdomen
- For stage IV patients with distant metastasis in the abdomen, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the abdomen ([dmets_abdomen] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Abdomen

[dx_to_dmets_abdomen_days], [dx_to_dmets_abdomen_mos], [dx_to_dmets_abdomen_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_abdomen_days]; months [dx_to_dmets_abdomen_mos]; or years [dx_to_dmets_abdomen_yrs] from diagnosis to date of first distant metastasis in the abdomen
- For stage IV patients with distant metastasis in the abdomen at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the abdomen ([dmets_abdomen] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Bone

[dmets_bone]

Value (Numeric)

- 0 = No distant metastasis in the bone
- 1 = Distant metastasis in the bone

Description

- Indicates the presence of distant metastasis in the bone post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the bone post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the bone at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Bone

[dx_to_dmets_bone_days], [dx_to_dmets_bone_mos], [dx_to_dmets_bone_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_bone_days]; months [dx_to_dmets_bone_mos]; or years [dx_to_dmets_bone_yrs] from diagnosis to date of first distant metastasis in the bone
- For stage IV patients with distant metastasis in the bone at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the bone ([dmets_bone] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Brain

[dmets_brain]

Value (Numeric)

- 0 = No distant metastasis in the brain
- 1 = Distant metastasis in the brain

Description

- Indicates the presence of distant metastasis in the brain post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the brain post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the brain at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Brain

[dx_to_dmets_brain_days], [dx_to_dmets_brain_mos], [dx_to_dmets_brain_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_brain_days]; months [dx_to_dmets_brain_mos]; or years [dx_to_dmets_brain_yrs] from diagnosis to date of first distant metastasis in the brain
- For stage IV patients with distant metastasis in the brain at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the brain ([dmets_brain] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Breast

[dmets_breast]

Value (Numeric)

- 0 = No distant metastasis in the breast
- 1 = Distant metastasis in the breast

Description

- Indicates the presence of distant metastasis in the breast post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the breast post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the breast at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Breast

[dx_to_dmets_breast_days], [dx_to_dmets_breast_mos], [dx_to_dmets_breast_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_breast_days]; months [dx_to_dmets_breast_mos]; or years [dx_to_dmets_breast_yrs] from diagnosis to date of first distant metastasis in the breast
- For stage IV patients with distant metastasis in the breast at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the breast ([dmets_breast] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in an Extremity

[dmets_extremity]

Value (Numeric)

- 0 = No distant metastasis in an extremity
- 1 = Distant metastasis in an extremity

Description

- Indicates the presence of distant metastasis in an extremity post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in an extremity post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in an extremity at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in an Extremity

[dx_to_dmets_extremity_days], [dx_to_dmets_extremity_mos], [dx_to_dmets_extremity_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_extremity_days]; months [dx_to_dmets_extremity_mos]; or years [dx_to_dmets_extremity_yrs] from diagnosis to date of first distant metastasis in an extremity
- For stage IV patients with distant metastasis in an extremity at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in an extremity ([dmets_extremity] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Head and Neck

[dmets_head_neck]

Value (Numeric)

- 0 = No distant metastasis in the head and neck
- 1 = Distant metastasis in the head and neck

Description

- Indicates the presence of distant metastasis in the head and neck post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the head and neck post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the head and neck at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Head and Neck

[dx_to_dmets_head_neck_days], [dx_to_dmets_head_neck_mos], [dx_to_dmets_head_neck_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_head_neck_days]; months [dx_to_dmets_head_neck_mos]; or years [dx_to_dmets_head_neck_yrs] from diagnosis to date of first distant metastasis in the head and neck
- For stage IV patients with distant metastasis in the head and neck at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in the head and neck ([dmets_head_neck] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Liver

[dmets_liver]

Value (Numeric)

- 0 = No distant metastasis in liver
- 1 = Distant metastasis in liver

Description

- Indicates the presence of distant metastasis in the liver post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the liver post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the liver at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Liver

[dx_to_dmets_liver_days], [dx_to_dmets_liver_mos], [dx_to_dmets_liver_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_liver_days]; months [dx_to_dmets_liver_mos]; or years [dx_to_dmets_liver_yrs] from diagnosis to date of first distant metastasis in liver
- For stage IV patients with distant metastasis in liver at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in liver ([dmets_liver] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Pelvis

[dmets_pelvis]

Value (Numeric)

- 0 = No distant metastasis in pelvis
- 1 = Distant metastasis in pelvis

Description

- Indicates the presence of distant metastasis in the pelvis post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the pelvis post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the pelvis at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Pelvis

[dx_to_dmets_pelvis_days], [dx_to_dmets_pelvis_mos], [dx_to_dmets_pelvis_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_pelvis_days]; months [dx_to_dmets_pelvis_mos]; or years [dx_to_dmets_pelvis_yrs] from diagnosis to date of first distant metastasis in pelvis
- For stage IV patients with distant metastasis in pelvis at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in pelvis ([dmets_pelvis] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Distant Metastasis in Thorax

[dmets_thorax]

Value (Numeric)

- 0 = No distant metastasis in thorax
- 1 = Distant metastasis in thorax

Description

- Indicates the presence of distant metastasis in the thorax post-diagnosis according to a radiology or pathology report (Appendix Table 5)
- For stage I-III patients, indicates the presence of distant metastasis in the thorax post-diagnosis
- For stage IV patients, indicates the presence of distant metastasis in the thorax at the time of diagnosis or post-diagnosis
- Populated only if cancer diagnosis is the patient's first BPC Project Cancer

Time (Days, Months, Years) from Diagnosis to Distant Metastasis in Thorax

[dx_to_dmets_thorax_days], [dx_to_dmets_thorax_mos], [dx_to_dmets_thorax_yrs]

Value (Numeric)

Description

- Interval in days [dx_to_dmets_thorax_days]; months [dx_to_dmets_thorax_mos]; or years [dx_to_dmets_thorax_yrs] from diagnosis to date of first distant metastasis in thorax
- For stage IV patients with distant metastasis in thorax at diagnosis, time in days, months, and years will be 0.
- Populated only among patients with distant metastasis in thorax ([dmets_thorax] = 1) and if cancer diagnosis is the patient's first BPC Project Cancer
- Note that this variable cannot be used as a time-to-event endpoint as it does not account for death or censoring.

Patient Received a Cancer-Directed Regimen Prior to Distant Metastasis

[reg_rcvd_before_distant_mets]

Value (Character)

- No
- Yes

Description

- Indicates whether the patient received a cancer-directed regimen prior to distant metastasis
- Populated only if diagnosed with distant metastasis ([dmets_stage_i_iii] = 1)

Overall Survival from Diagnosis: Status Indicator

[os_dx_status]

Value (Numeric)

- 1 = Dead
- 0 = Censored

Description

- An event is defined by death
- Patients were censored if not known to be dead

Overall Survival from Diagnosis (Days, Months, Years)

[tt_os_dx_days], [tt_os_dx_mos], [tt_os_dx_yrs]

Value (Numeric)

Description

- Time from diagnosis to death or the date last known alive
- Interval in days [tt_os_dx_days]; months [tt_os_dx_mos]; or years [tt_os_dx_yrs] from cancer diagnosis to last known alive date or death

Overall Survival from Advanced Disease: Status Indicator

[os_adv_status]

Value (Numeric)

- 1 = Dead
- 0 = Censored

Description

- An event is defined by death
- Patients were censored if not known to be dead
- Populated only if:
    - BPC Project Cancer
    - Patient diagnosed with stage IV cancer ([stage_dx] = Stage IV) or diagnosed with a distant metastasis ([dmets_stage_i_iii] = 1)

Overall Survival from Advanced Disease (Days, Months, Years)

[tt_os_adv_days], [tt_os_adv_mos], [tt_os_adv_yrs]

Value (Numeric)

Description

- Time from advanced disease to death or the date last known alive
- Interval in days [tt_os_adv_days]; months [tt_os_adv_mos]; or years [tt_os_adv_yrs] from diagnosis of advanced disease to last known alive date or death
- Populated only for BPC Project Cancers that were stage IV ([stage_dx] = Stage IV) or diagnosed with a distant metastasis ([dmets_stage_i_iii] = 1)

Progression-Free Survival Cohort

[pfs_cohort]

Value (Character)

- Stage I-III with Distant Mets
- Stage IV

Description

- Indicates whether PFS was defined from first date of distant metastasis (stage I-III) or from diagnosis (stage IV)
- Populated only if BPC Project Cancer

Progression Free Survival-Imaging (PFS-I): Status Indicator

[pfs_i_adv_status]

Value (Numeric)

- 1 = Progression according to radiologist assessment; death
- 0 = Censored

Description

- An event is defined by:
    - Radiologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Death
- Patients were censored in the absence of an event
- Populated only for BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Imaging (PFS-I)

[tt_pfs_i_adv_days], [tt_pfs_i_adv_mos], [tt_pfs_i_adv_yrs]

Value (Numeric)

Description

- Interval in days [tt_pfs_i_adv_days]; months [tt_pfs_i_adv_mos]; or years [tt_pfs_i_adv_yrs] from stage IV diagnosis or date of distant metastasis (stage I-III at diagnosis) to:
    - Radiologist assessment of progressing/worsening/enlarging
    - Death
    - Last known alive date, if patient is censored according to PFS-I status indicator [pfs_i_adv_status]
- Populated only for BPC Project Cancers that were stage IV at diagnosis or stage I-III at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival - Medical Oncologist Assessment (PFS-M): Status Indicator

[pfs_m_adv_status]

Value (Numeric)

- 1 = Progression according to medical oncologist assessment; death
- 0 = Censored

Description

- An event is defined by:
    - Medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Death
- Patients were censored in the absence of an event
- Populated only for BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Medical Oncologist (PFS-M)

[tt_pfs_m_adv_days], [tt_pfs_m_adv_mos], [tt_pfs_m_adv_yrs]

Value (Numeric)

Description

- Interval in days [tt_pfs_m_adv_days]; months [tt_pfs_m_adv_mos]; or years [tt_pfs_m_adv_yrs] from stage IV diagnosis or date of distant metastasis (stage I-III at diagnosis) to:
    - Medical oncologist assessment of progressing/worsening/enlarging
    - Death
    - Last known alive date, if patient is censored according to PSF-M status indicator [pfs_m_adv _status]
- Populated only for BPC Project Cancers that were stage IV at diagnosis or stage I-III at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Imaging or Medical Oncologist Assessment (PFS-I-or-M): Status Indicator

[pfs_i_or_m_adv_status]

Value (Numeric)

- 1 = Progression according to radiologist assessment or medical oncologist assessment; death
- 0 = Censored

Description

- An event is defined by:
    - Radiologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Death
- Patients were censored in the absence of an event
- Populated only for BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Imaging or Medical Oncologist Assessment (PFS-I-or-M)

[tt_pfs_i_or_m_adv_days], [tt_pfs_i_or_m_adv_mos], [tt_pfs_i_or_m_adv_yrs]

Value (Numeric)

Description

- Time from diagnosis to radiologist impression of progressing/worsening/enlarging cancer status, last known alive date, or death
- Interval in days [tt_pfs_i_or_m_adv_days]; months [tt_pfs_i_or_m_adv_mos]; or years [tt_pfs_i_or_m_adv_yrs] from stage IV diagnosis or date of distant metastasis (stage I-III at diagnosis) to:
    - The first of a radiologist assessment of progressing/worsening/enlarging, medical oncologist assessment of progressing/worsening/enlarging, or death
    - Last known alive date, if patient is censored according to PFS-I-or-M status indicator [pfs_i_or_m_adv_status]
- Populated only for BPC Project Cancers that were stage IV at diagnosis or stage I-III at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Imaging and Medical Oncologist Assessment (PFS-I-and-M): Status Indicator

[pfs_i_and_m_adv_status]

Value (Numeric)

- 1 = Progression according to radiologist assessment and medical oncologist assessment; death
- 0 = Censored

Description

- An event is defined by:
    - Radiologist assessment and medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Death
- Patients were censored in the absence of an event
- Populated only for BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1).

Progression Free Survival-Imaging and Medical Oncologist Assessment (PFS-I-and-M)

[tt_pfs_i_and_m_adv_days], [tt_pfs_i_and_m_adv_mos], [tt_pfs_i_and_m_adv_yrs]

Value (Numeric)

Description

• Time from diagnosis to radiologist assessment or medical oncologist assessment of progressing/worsening/enlarging cancer status, last known alive date or death
• Interval in days [tt_pfs_i_and_m_adv_days]; months [tt_pfs_i_and_m_adv_mos]; or years [tt_pfs_i_and_m_adv_yrs] from stage IV diagnosis or date of distant metastasis (stage I-III at diagnosis) to:
    • The latter of radiologist assessment or medical oncologist assessment documenting progressing disease, if the patient progressed according to both sources
    • Death
    • Last known alive date, if patient is censored according to PFS-I-and-M status indicator [pfs_i_and_m_adv _status]
• Populated only for BPC Project Cancers that were stage IV at diagnosis or stage I-III at diagnosis with subsequent distant metastasis [dmets_stage_i_iii = 1].

Release Version Number

[release_version]

Value (Character)

• 2.0-public

Description

• Indicates the version number of the data release

CANCER-DIRECTED REGIMEN DATASET

The Cancer-Directed Regimen dataset is structured as one record per regimen-associated cancer diagnosis, per patient. For example, if a regimen is associated with a single cancer diagnosis, there will be one corresponding record in this dataset. If a regimen is associated with two cancer diagnoses, then there will be two corresponding records in this dataset: one for the first associated cancer diagnosis and another for the second associated cancer diagnosis. If it is unknown which cancer diagnosis the regimen is associated with, there will still be one record in this dataset.

Cancer-directed regimens were curated for all cancer diagnoses, including both BPC Project and non-BPC Project Cancers. A regimen can only consist of one drug or up to five drugs given together. Cancer-directed drugs include anti-neoplastic drugs, immunotherapies, targeted therapies, and hormone therapies. A break in treatment of ≥8 weeks was used to indicate the end of a regimen; even if all drugs in the regimen were re-initiated 8+ weeks later, this was considered a new regimen. All LHRH agonists ([regimen_drugs] containing "Goserlin Acetate", "Histrelin Acetate", "Leuprolide Acetate", "Triptorelin") were curated as separate regimens even if administered with additional drugs.

Compliance with data privacy requires redaction of the name and duration of investigational drugs. If the cancer-directed drug was part of an investigational drug trial, the drug name(s) [drugs_drug_1-drugs_drug_5] were set to "Investigational Drug" and the end date interval will match the start date interval. Identification of investigational drugs varies by institution depending on contractual obligations. Each institution followed their respective data privacy specifications for investigational drug masking.

This dataset can be linked to the following datasets:

- *BPC Project Cancer Diagnosis, Non-BPC Project Cancer Diagnosis, Cancer Panel Test datasets using the variables [cohort], [record_id] and [ca_seq]*
- *Patient Characteristics, PRISSMM Pathology, PRISSMM Imaging, and PRISSMM Medical Oncologist Assessment datasets using [cohort] and [record_id].*

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The recommended variable is noted in the description.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, "GENIE"; the second component is the institution's abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Cancer Sequence Associated with Regimen

[ca_seq]

Value (Numeric)

- 0 = first and only cancer
- 1 = first of two or more primaries
- 2 = second of two or more primaries
- …10 = tenth of ten or more primaries

Description

- The cancer sequence number associated with this regimen
- When regimens are associated with multiple cancer diagnoses, each associated cancer diagnosis will be a separate row in the dataset
- Populated only if it is known which diagnosis is associated with this cancer-directed regimen
- Primary key for the Cancer Diagnosis, Cancer-Directed Regimen and Cancer Panel Test datasets

Regimen Number

[regimen_number]

Value (Numeric)

- 1 = first regimen
- 2 = second regimen
- …n = nth regimen

Description

- Order for cancer-directed regimens based on the start date of the first cancer-directed drug
- The cancer-directed regimen dataset is structured as one row per regimen-associated cancer diagnosis, per patient. For example, if a regimen is associated with a single cancer diagnosis, there will be one corresponding record in this dataset. If a regimen is associated with two cancer diagnoses, then there will be two corresponding records in this dataset: one for the first associated cancer diagnosis and another for the second associated cancer diagnosis. If it is unknown which diagnosis the regimen is associated with, there will be one row in the cancer-directed regimen dataset.
    - This implies that multiple records will have the same regimen number in this dataset when they are associated with multiple cancer diagnoses (e.g., if the first cancer-directed regimen that a patient received is associated with their first and second cancer diagnosis then there will be two rows in this dataset, one for the first cancer diagnosis and one for the second cancer diagnosis, and both will have [regimen_number] = 1).

Regimen Number Within Cancer Diagnosis

[regimen_number_within_cancer]

Value (Numeric)

- 1 = first regimen associated with specific cancer diagnosis
- 2 = second regimen associated with specific cancer diagnosis
- …n = nth regimen associated with specific cancer diagnosis

Description

- Order of cancer-directed regimens that are associated with a specific cancer diagnosis [ca_seq] based on the start date of the first cancer-directed drug

BPC Project (Index) Cancer Indicator

[redcap_ca_index]

Value (Character)

- Yes
- No

Description

- Indicates whether the regimen was associated with a BPC Project Cancer
- The BPC Project Cancer is defined as the cancer that met eligibility criteria, underwent genomic sequencing and was submitted to AACR Project GENIE.
- Each patient has at least one BPC Project Cancer. Patients may have multiple BPC Project Cancers, though this is rare.
- PRISSMM data elements are curated for BPC Project Cancers.
- The terms "BPC Project Cancer" and "index cancer" are used interchangeably.
- Further details regarding the definition of BPC Project and non-BPC Project Cancers can be found in Appendix 1. BPC Project and Non-BPC Project Cancers.

Data Standard:
- ICD-O-3 topography code

Number of Cancer-Directed Drugs in a Regimen

[drugs_num]

Value (Numeric)

- 1-5

Description

- Number of cancer-directed drugs in a regimen; up to 5 recorded

Institution That Administered/Ordered Regimen

[drugs_inst]

Value (Character)

- At the internal/native institution only
- Split across internal and external institution
- At external institution only

Description

- Location where regimen was administered/ordered
- Indicates whether regimen was administered/ordered at the internal (i.e. the same institution as the [institution] variable) or external institution.

Institution That First Ordered the Regimen

[drugs_firstinst]

Value (Character)

- Internal institution
- External institution

Description

- Location where regimen was first ordered
- Indicates whether regimen was administered/ordered at the internal (i.e. the same institution as the [institution] variable) or external institution.
- Populated only if the regimen was administered/ordered across internal and external institutions ([drugs_inst] = “Split across internal and external institution”)

Regimen Was Part of a Clinical Trial

[drugs_ct_yn]

Value (Character)

- Yes
- No

Description

- Indicator for whether a cancer-directed drug was part of a clinical trial
- If the cancer-directed drug was part of an investigational drug trial, the drug name(s) [drugs_drug_1-drugs_drug_5] will be masked as “Investigational drug” and the end date interval will match the start date interval.
    - Identification of investigational drugs varies by institution depending on contractual obligations. For some institutions, all drugs that are part of an investigational trial are required to be masked, even if standard of care. For other institutions, only the investigational drug(s) are required to be masked.

Regimen Discontinuation Status

[drugs_dc_ynu]

Value (Character)

- Yes
- No
- Unknown, no documentation found

Description

• Indicator for whether the drug regimen was discontinued. If not discontinued, the patient was still receiving cancer-directed regimen at time of curation, or it is unknown whether the regimen had ended
• The response to this variable specifies whether there will be a cancer-directed regimen end date or a date of last administration specified in the variables [dx_drug_end_or_lastadm_int_1 – dx_drug_end_or_lastadm_int_5]

Names of Drugs in Regimen

[regimen_drugs]

Value (Character)

Description

• Names of cancer-directed drugs received together
• Concatenation of variables [drugs_drug_1]-[drugs_drug_5]

Name of Cancer-Directed Drug in Regimen, Drugs 1-5

[drugs_drug_1] – [drugs_drug_5]

Value (Character)

• Name of cancer-directed drug 1 through 5 in each regimen (Appendix 2)

Description

• The cancer drug label contains the generic/ingredient name with the synonyms in parentheses (e.g. Nivolumab (BMS936558, MDX1106, NIVO, ONO4538, Opdivo))
• If the cancer-directed drug was part of an investigational drug trial, the drug name(s) [drugs_drug_1-drugs_drug_5] will be masked as "Investigational Drug" and the end date interval will match the start date interval.
    • Identification of investigational drugs varies by institution depending on contractual obligations. For some institutions, all drugs that are part of an investigational trial are required to be masked, even if standard of care. For other institutions, only the investigational drug(s) are required to be masked.

Data Standard: National Cancer Institute Thesaurus: Antineoplastic Agents

Time (Days) from Date of Birth to Start of Cancer-Directed Drug in Regimen, Drugs 1-5

[drugs_startdt_int_1] – [drugs_startdt_int_5]

Value (Numeric)

Description

- Interval in days from date of birth to start of each cancer-directed drug 1 – drug 5

Time (Days, Months) from Associated Cancer Diagnosis to Start of Cancer-Directed Drug in Regimen, Drugs 1-5

[dx_drug_start_int_1] – [dx_drug_start_int_5], [dx_drug_start_int_mos_1] – [dx_drug_start_int_mos_5]

Value (Numeric)

Description

- Interval in days [dx_drug_start_int_1] – [dx_drug_start_int_5] or months [dx_drug_start_int_mos_1] – [dx_drug_start_int_mos_5] from cancer diagnosis to start of cancer-directed drug 1 – drug 5

Time (Days, Months) from Associated Cancer Diagnosis to End of Cancer-Directed Drug in Regimen, Drugs 1-5

[dx_drug_end_int_1] – [dx_drug_end_int_5], [dx_drug_end_int_mos_1] – [dx_drug_end_int_mos_5]

Value (Numeric)

Description

- Number of days ([dx_drug_end_int_1-dx_drug_end_int_5]) or months ([dx_drug_end_int_mos_1-dx_drug_end_int_mos5]) from cancer diagnosis to end of cancer-directed drug 1 – drug 5
- If the cancer-directed drug was part of an investigational drug trial, the drug name(s) [drugs_drug_1-drugs_drug_5] will be set to "Investigational drug" and the end date interval will match the start date interval.
    - Identification of investigational drugs varies by institution depending on contractual obligations. For some institutions, all drugs that are part of an investigational trial are required to be masked, even if standard of care. For other institutions, only the investigational drug(s) are required to be masked.
- Populated only if drug regimen was discontinued ([drugs_dc_ynu]="Yes")

Time (Days) from Associated Cancer Diagnosis to End Date or Last Known Administration Date of Drugs 1-5

[dx_drug_end_or_lastadm_int_1]-[dx_drug_end_or_lastadm_int_5]

Value (Numeric)

Description

- Number of days from associated cancer diagnosis to the end of cancer-directed drug 1 – drug 5 if regimen was discontinued ([drugs_dc_ynu] = "Yes") or to last known administration date of cancer-directed drug 1 – drug 5 if regimen was not discontinued ([drugs_dc_ynu] = "No" or "Unknown")

Time from Associated Cancer Diagnosis to Start of Cancer-Directed Regimen

[dx_reg_start_int], [dx_reg_start_int_mos], [dx_reg_start_int_yrs]

Value (Numeric)

Description

- Interval in days [dx_reg_start_int]; months [dx_reg_start_int_mos]; or years [dx_reg_start_int_yrs] from associated cancer diagnosis to the start of the first drug in cancer-directed regimen

Time (Days, Months, Years) from Associated Cancer Diagnosis to End of First Drug Discontinued in Cancer-Directed Regimen

[dx_reg_end_any_int], [dx_reg_end_any_int_mos], [dx_reg_end_any_int_yrs]

Value (Numeric)

Description

- Number of days [dx_reg_end_any_int]; months [dx_reg_end_any_int_mos]; or years [dx_reg_end_any_int_yrs] from associated cancer diagnosis to the end of first drug discontinued in cancer-directed regimen
- Populated only if regimen is known to be discontinued ([drugs_dc_ynu] = “Yes”)

Time (Days, Months, Years) from Associated Cancer Diagnosis to End of All Drugs in Cancer-Directed Regimen

[dx_reg_end_all_int], [dx_reg_end_all_int_mos], [dx_reg_end_all_int_yrs]

Value (Numeric)

Description

- Number of days [dx_reg_end_all_int]; months [dx_reg_end_all_int_mos]; or years [dx_reg_end_all_int_yrs] from associated cancer diagnosis to the end of all drugs in the cancer-directed regimen
- Populated only if regimen is known to be discontinued ([drugs_dc_ynu] = “Yes”)

Overall Survival from Start of Cancer-Directed Drug: Status Indicator

[os_d_status]

Value (Numeric)

- 1 = Dead
- 0 = Censored

Description

- An event is defined by death
- Patients were censored if not known to be dead

Overall Survival from Start of Cancer-Directed Drug (Days, Months, Years)

[tt_os_d1_days] – [tt_os_d5_days], [tt_os_d1_mos] – [tt_os_d5_mos], [tt_os_d1_yrs] – [tt_os_d5_yrs]

Value (Numeric)

Description

- Time from start of cancer-directed drug to death/last known alive date
- Interval in days [tt_os_d1_days-tt_os_d5_days]; months [tt_os_d1_mos-tt_os_d5_mos]; or years [tt_os_d1_yrs-tt_os_d5_yrs] from start of cancer-directed drug [drugs_startdt_int_drug_1-drugs_startdt_int_drug_5] to death [hybrid_death_int] or last known alive date [dob_lastalive_int]

Overall Survival from Start of Cancer-Directed Regimen: Status Indicator

[os_g_status]

Value (Numeric)

- 1 = Dead
- 0 = Censored

Description

- An event is defined by death
- Patients were censored if not known to be dead

Overall Survival from Start of Cancer-Directed Regimen (Days, Months, Years)

[tt_os_g_days], [tt_os_g_mos], [tt_os_g_yrs]

Value (Numeric)

Description

- Time from start of cancer-directed regimen to death/last known alive date
- Interval in days [tt_os_g_days]; months [tt_os_g_mos]; or years [tt_os_g_yrs] from the start of the cancer-directed regimen to death [hybrid_death_int] or last known alive date [dob_lastalive_int]

Progression Free Survival – Imaging (PFS-I) from Start of Cancer-Directed Regimen: Status Indicator

[pfs_i_g_status]

Value (Numeric)

- 1 = Progression or death
- 0 = Censored

Description

- An event is defined by:
    - Radiologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    - Death
- Patients were censored at the time of next cancer-directed regimen or end of follow up.
    - Since LHRH agonists were recorded as separate regimens, the next cancer-directed regimen may begin before the LHRH agonist treatment ends.
- Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis [dmets_stage_i_iii = 1].

Progression Free Survival – Imaging (PFS-I) from Start of Cancer-Directed Regimen

[tt_pfs_i_g_days], [tt_pfs_i_g_mos], [tt_pfs_i_g_yrs]

Value (Numeric)

Description

• Interval in days [tt_pfs_i_g_days]; months [tt_pfs_i_g_mos]; or years [tt_pfs_i_g_yrs] from start of cancer-directed regimen to:
    • Radiologist assessment of progressing/worsening/enlarging
    • Death
    • Start of next regimen (if applicable) or last known alive date, if patient is censored according to PFS-I status indicator [pfs_i_g_status]
• Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1). For patients with stage I-III disease at diagnosis with subsequent distant metastasis, PFS is only defined for regimens that began after date of distant metastasis

Progression Free Survival – Medical Oncologist Assessment (PFS-M) from Start of Cancer-Directed Regimen: Status Indicator

[pfs_m_g_status]

Value (Numeric)

• 1 = Progression according to medical oncologist assessment; death
• 0 = Censored

Description

• An event is defined by:
    • Medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    • Death
• Patients were censored at the time of next cancer-directed regimen or end of follow up.
    • Since LHRH agonists were recorded as separate regimens, the next cancer-directed regimen may begin before the LHRH agonist treatment ends.
• Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1). For patients with stage I-III disease at diagnosis with subsequent distant metastasis, PFS is only defined for regimens that began after date of distant metastasis

Progression Free Survival – Medical Oncologist Assessment (PFS-M) from Start of Cancer-Directed Regimen

[tt_pfs_m_g_days], [tt_pfs_m_g_mos], [tt_pfs_m_g_yrs]

Value (Numeric)

Description

• Interval in days [tt_pfs_m_g_days]; months [tt_pfs_m_g_mos]; or years [tt_pfs_m_g_yrs] from start of cancer-directed regimen to:
    • Medical oncologist assessment of mixed or progressing disease
    • Death
    • Start of next regimen (if applicable) or last known alive date, if patient is censored according to PFS-M status indicator [pfs_m_g_status]
• Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1). For patients with stage I-III disease at diagnosis with subsequent distant metastasis, PFS is only defined for regimens that began after date of distant metastasis.

Progression Free Survival – Imaging or Medical Oncologist Assessment (PFS-I-or-M) from Start of Cancer-Directed Regimen: Status Indicator

[pfs_i_or_m_g_status]

Value (Numeric)

• 1 = Progression according to radiologist assessment or medical oncologist assessment; death
• 0 = Censored

Description

• An event is defined by:
    • Radiologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    • Medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    • Death
• Patients were censored at the time of next cancer-directed regimen or end of follow up.
    • Since LHRH agonists were recorded as separate regimens, the next cancer-directed regimen may begin before the LHRH agonist treatment ends.
• Populated only for cancer-directed regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis [dmets_stage_i_iii = 1].

Progression Free Survival-Imaging or Medical Oncologist Assessment (PFS-I-or-M)

[tt_pfs_i_or_m_g_days], [tt_pfs_i_or_m_g_mos], [tt_pfs_i_or_m_g_yrs]

Value (Numeric)

Description

• Interval in days [tt_pfs_i_or_m_g_days]; months [tt_pfs_i_or_m_g_mos]; or years [tt_pfs_i_or_m_g_yrs] from start of cancer-directed regimen to:
    • The first of a radiologist assessment of progressing/worsening/enlarging, medical oncologist assessment of progressing/worsening/enlarging, or death
    • Start of next regimen (if applicable) or last known alive date, if patient is censored according to PFS-I-or-M status indicator [pfs_i_or_m_g_status]
• Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1). For patients with stage I-III disease at diagnosis with subsequent distant metastasis, PFS is only defined for regimens that began after date of distant metastasis.

Progression Free Survival – Imaging and Medical Oncologist Assessment (PFS-I-and-M) from Start of Cancer-Directed Regimen: Status Indicator

[pfs_i_and_m_g_status]

Value (Numeric)

• 1 = Progression according to radiologist assessment and medical oncologist assessment; death
• 0 = Censored

Description

• An event is defined by:
    • Radiologist assessment and medical oncologist assessment of the change in the patient's cancer status as progressing/worsening/enlarging
    • Death
• Patients were censored at the time of next cancer-directed regimen or end of follow up.
    • Since LHRH agonists were recorded as separate regimens, the next cancer-directed regimen may begin before the LHRH agonist treatment ends.
• Populated only for cancer-directed regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis [dmets_stage_i_iii = 1].

Progression Free Survival-Imaging and Medical Oncologist Assessment (PFS-I-and-M)

[tt_pfs_i_and_m_g_days], [tt_pfs_i_and_m_g_mos], [tt_pfs_i_and_m_g_yrs]

Value (Numeric)

Description

• Interval in days [tt_pfs_i_and_m_g_days]; months [tt_pfs_i_and_m_g_mos]; or years [tt_pfs_i_and_m_g_yrs] from start of cancer-directed regimen to:
    • The latter of a radiologist assessment of progressing/worsening/enlarging, medical oncologist assessment of progressing/worsening/enlarging, or death
    • Start of next regimen (if applicable) or last known alive date, if patient is censored according to PFS-I-and-M status indicator [pfs_i_and_m_g_status]
• Populated only for regimens associated with BPC Project Cancers with stage IV disease at diagnosis or stage I-III disease at diagnosis with subsequent distant metastasis ([dmets_stage_i_iii] = 1). For patients with stage I-III disease at diagnosis with subsequent distant metastasis, PFS is only defined for regimens that began after date of distant metastasis.

Time (Days, Months, Years) to Next Treatment For Any Cancer: Status Indicator

[ttnt_any_ca_status]

Value (Numeric)

• 1 = Initiation of a subsequent cancer-directed regimen for any cancer; death
• 0 = Censored

Description

• An event is defined by:
    • Initiation of a subsequent cancer-directed regimen for any cancer diagnosis
    • Death
• Patients were censored in the absence of an event

Time (Days, Months, Years) to Next Treatment for Any Cancer: Status Indicator

[ttnt_any_ca_days], [ttnt_any_ca_mos], [ttnt_any_ca_yrs]

Value (Numeric)

Description

• Time from start of cancer-directed regimen to initiation of subsequent cancer-directed regimen for any cancer or death
• Interval in days [ttnt_any_ca_days]; months [ttnt_any_ca_mos]; or years [ttnt_any_ca_yrs] from start of cancer-directed regimen to:
    • Initiation of subsequent cancer-directed regimen for any cancer
    • Death
    • Last known alive date, if patient is censored according to time to next treatment (any cancer) status indicator [ttnt_any_ca_days]
• Note: TTNT is not defined for LHRH agonists ([regimen_drugs] containing “Goserlin Acetate”, “Histrelin Acetate”, “Leuprolide Acetate”, “Triptorelin”), and they are excluded from TTNT calculations for other regimens

Time to Next Treatment For This Cancer: Status Indicator

[ttnt_ca_seq_status]

Value (Numeric)

- 1 = Initiation of a subsequent cancer-directed regimen for the same cancer; death
- 0 = Censored

Description

- An event is defined by:
  - Initiation of a subsequent cancer-directed regimen for the same cancer diagnosis
  - Death
- Patients were censored in the absence of an event

Time (Days, Months, Years) to Next Treatment for this Cancer: Status Indicator

[ttnt_ca_seq_days], [ttnt_ca_seq_mos], [ttnt_ca_seq_yrs]

Value (Numeric)

Description

- Time from start of cancer-directed regimen to initiation of subsequent cancer-directed regimen for the same cancer or death
- Interval in days [ttnt_ca_seq_days]; months [ttnt_ca_seq_mos]; or years [ttnt_ca_seq_yrs] from start of cancer-directed regimen to:
  - Initiation of subsequent cancer-directed regimen for this cancer
  - Death
  - Last known alive date, if patient is censored according to time to next treatment (the same cancer) status indicator [ttnt_ca_seq_days]
- Note: TTNT is not defined for LHRH agonists ([regimen_drugs] containing “Goserlin Acetate”, “Histrelin Acetate”, “Leuprolide Acetate”, “Triptorelin”), and they are excluded from TTNT calculations for other regimens

Release Version Number

[release_version]

Value (Character)

- 2.0-public

Description

- Indicates the version number of the data release

PRISSMM PATHOLOGY DATASET

The Pathology dataset is structured as one record per pathology report, per patient. All pathology reports beginning with the month and year of the first BPC Project Cancer diagnosis are curated. All subsequent pathology reports were recorded (including pathology reports corresponding to non-BPC Project Cancer and subsequent BPC Project Cancer diagnoses; Appendix 1). Additionally, all non-BPC Project NSCLC pathology reports are curated.

The PRISSMM Pathology dataset can be linked to the following datasets:

- *Cancer Panel Test dataset using [cohort], [record_id], [path_proc_number] and [path_report_number]*
- *Patient Characteristics, BPC Project and non-BPC Project Cancer Diagnosis, Cancer-Directed Regimen, PRISSMM Imaging, and PRISSMM Medical Oncologist Assessment datasets using [cohort] and [record_id].*

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The preferred variable is noted in the description.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, “GENIE”; the second component is the institution’s abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Pathology Procedure Number

[path_proc_number]

Value (Numeric)

Description

- Order of pathology procedures based on the date of the procedure [path_proc_int]
- Pathology procedures occurring on the same date have the same procedure number
- Primary key for PRISSMM pathology and cancer panel test datasets.

Pathology Report Number

[path_rep_number]

Value (Numeric)

Description

- Order of pathology reports from the same pathology procedure [path_proc_number]
- For example, a pathology procedure with three associated reports will have pathology report numbers 1-3.
- Primary key for PRISSMM pathology and cancer panel test datasets.

Institution Where Procedure Was Performed

[path_proc_inst]

Value (Character)

- Internal institution
- External institution

Description

- Indicates whether the pathology procedure was performed at the internal (i.e. the same institution as the [institution] variable) or external institution.

Institution Where Pathology was Reviewed

[path_rep_inst]

Value (Character)

- Internal institution
- External institution

Description

- Indicates whether the pathology procedure was reviewed at the internal (i.e. the same institution as the [institution] variable) or external institution.

Time (Days) from Date of Birth to Pathology Procedure Date

[path_proc_int]

Value (Numeric)

Description

- Interval in days from date of birth to pathology procedure date

Time (Days, Months, Years) from First BPC Project Cancer to Pathology Procedure Date

[dx_path_proc_days], [dx_path_proc_mos], [dx_path_proc_yrs]

Value (Numeric)

Description

- Interval in days [dx_path_proc_days]; months [dx_path_proc_mos]; or years [dx_path_proc_yrs] from first BPC Project Cancer diagnosis to pathology procedure date

Pathology Type

[path_proc_type]

Value (Character)

- Cytology
- Surgical pathology
- Other

Description

- Type of pathology

Number of Specimens Included in the Pathology Report

[path_num_spec]

Value (Numeric)

- 1-25

Description

- The number of distinct specimens from one procedure that are included in the report
- Up to 25 specimens can be curated from each pathology report

Anatomic Site for Each Specimen 1-25

[path_site1] – [path_site25]

Value (Character)

- ICD-O-3 topography code

Description

- The anatomic site of the specimen is often different than the type of invasive cancer

In Situ Cancer Found in at Least One Specimen in the Pathology Report

[path_insitu_any]

Value (Character)

- Yes
- No

Description

- Indicates whether in situ cancer was found in any of the 25 specimens on the pathology report
- Based on [path_insitu1] – [path_insitu25]

Number of Specimens with In Situ Cancer in Pathology Report

[n_specimen_insitu]

Value (Numeric)

- 0-25

Description

- Number of specimens with in situ cancer on the pathology report
- Based on [path_insitu1] – [path_insitu25]

In Situ Cancer Identified in Specimen 1-25

[path_insitu1] – [path_insitu25]

Value (Character)

- Yes
- No

Description

- Indicates whether any in situ cancer (ICD-O-3 Behavior Code = 2) is present in the specimen

Invasive Cancer Found in at Least One Specimen in the Pathology Report

[path_ca_inv_any]

Value (Character)

- Yes
- No

Description

- Indicates whether invasive cancer was found in any of the 25 specimens on the pathology report
- Based on [path_ca1] – [path_ca25]

Number of Specimens with Invasive Cancer in Pathology Report

[n_specimen_inv]

Value (Numeric)

- 0-25

Description

- Number of specimens with invasive cancer in pathology report
- Based on [path_ca1] – [path_ca25]

Invasive Cancer Identified in Specimen 1-25

[path_ca1] – [path_ca25]

Value (Character)

- Yes
- No

Description

- Indicates whether any invasive cancer (ICD-O-3 Behavior Code = 3) is present in the specimen

Invasive Cancer Type for Each Specimen 1-25

[path_ca_type1] – [path_ca_type25]

Value (Character)

- Adrenocortical Carcinoma
- Anal Cancer
- Appendix Cancer
- Bile Duct Cancer
- Bladder Cancer
- Brain Cancer
- Breast Cancer
- NET or Carcinoid
- Cervical Cancer
- Colon Cancer
- Colon/Rectum Cancer
- Esophagus Cancer
- Ewing Sarcoma
- Fallopian Tube Cancer
- Gallbladder Cancer
- Germ Cell Tumor
- GIST
- Head and Neck Cancer
- Mesothelioma
- Ill Defined/Cancer of Unknown Primary
- Liver Cancer
- Lung Cancer, NOS
- Melanoma
- Merkel Cell
- Neuroblastoma
- Non Small Cell Lung Cancer
- Osteosarcoma
- Ovarian Cancer
- Pancreatic Cancer
- Parathyroid Cancer
- Penis Cancer
- Peritoneum Cancer
- Placenta Cancer
- Prostate Cancer

- Rectum and Rectosigmoid Cancer
- Renal Kidney Cancer
- Renal Pelvis Cancer
- Retinoblastoma
- Rhabdomyosarcoma
- Scrotum Cancer
- Small Cell Lung Cancer
- Small Intestine Cancer
- Stomach Cancer
- Testis Cancer
- Thymus Cancer
- Thyroid Cancer
- Uterus Cancer
- Vagina Cancer
- Vulva Cancer
- Wilms Tumor
- Other
- Not stated
- Unknown

Description

- Indicates cancer type identified in the specimen with invasive cancer
- Populated only if invasive cancer is present ([path_ca1] –[path_ca25] = “Yes”)

Cancer Histology Type for Specimen 1-25

[path_ca_hist1] – [path_ca_hist25]

Value (Character)

- ICD-O-3 topography code

Description

- Cancer histology type associated with specimen with invasive cancer
- Populated only if invasive cancer is present ([path_ca1] –[path_ca25] = “Yes”)

Biomarkers

Up to three tests for PD-L1 can be associated with a pathology report. The biomarker information corresponding to each of the three tests is consistent across variables, i.e. information corresponding to the first PD-L1 test is stored in variables [pdl1_yn], [pdl1_prepaint], [pdl1_test], etc. and the information corresponding to the second PD-L1 test is stored in variables [pdl1_yn_2], [pdl1_prepaint_2], [pdl1_test_2], etc. Summary variables indicating any testing and any positive result across all three instances are also provided (i.e. [pdl1_testing] and [pdl1_positive_any]).

Any PD-L1 Testing Reported on Pathology Report

[pdl1_testing]

Value (Character)

- Yes
- No

Description

- Indicates whether any PD-L1 testing was reported on a given pathology report
- Combines variables [pdl1_yn], [pdl1_yn_2], [pdl1_yn_3]

PD-L1 Testing Reported

[pdl1_yn], [pdl1_yn_2], [pdl1_yn_3]

Value (Character)

- Yes
- No

Description

- Indicates whether PD-L1 testing was reported on a given pathology report
- Up to three PD-L1 tests on a single pathology report are curated
- Populated only if any in situ or invasive cancer identified in specimens 1-25 ([path_insitu1]-[path_insitu25] = “Yes” or [path_ca1]-[path_ca25] = “Yes”)
- Preferred derived variable [pdl1_testing] incorporates all three PD-L1 tests.

Time (Days) from Date of Birth to PD-L1 Report Date

[pdl1_prepaint], [pdl1_prepaint_2], [pdl1_prepaint_3]

Value (Numeric)

Description

- Interval in days from date of birth to pathology report PD-L1 test result
- Populated only if PD-L1 testing reported ([pdl1_yn], [pdl1_yn_2], [pdl1_yn_3] = "Yes")

PD-L1 Antibody Test Type

[pdl1_test], [pdl1_test_2], [pdl1_test_3]

Value (Character)

- 22C3
- 28-2
- E1L3N
- SP142
- Other
- Unknown

Description

- PD-L1 test identified in pathology report
- Up to three tests can be recorded per pathology report
- Populated only if PD-L1 testing reported ([pdl1_yn], [pdl1_yn_2], [pdl1_yn_3] = "Yes")

PD-L1 Results Presented as a Percentage or Percentage Range of Tumor Cells

[pdl1_type___1], [pdl1_type_2___1], [pdl1_type_3___1]

Value (Character)

- Percentage or Percentage Range of Tumor Cells

Description

- Indicates that PD-L1 test results are represented as percentage or percentage range of tumor cells
- Populated only if PD-L1 results are represented as the percentage or percentage range of tumor cells

PD-L1 Results Presented as a Percentage or Percentage Range of Infiltrating Immune Cells

[pdl1_type___2], [pdl1_type_2___2], [pdl1_type_3___2]

Value (Character)

- Percentage or Percentage Range of Infiltrating Immune Cells

Description

- Indicates that PD-L1 test results are represented as the percentage or percentage range of infiltrating immune cells
- Populated only if PD-L1 results are represented as the percentage or percentage range of infiltrating immune cells

PD-L1 Results Presented as a Numeric Combined Positive Score

[pdl1_type___3], [pdl1_type_2___3], [pdl1_type_3___3]

Value (Character)

- Numeric (Combined Positive Score)

Description

- Indicates that PD-L1 test results are represented as a numeric combined positive score
- Populated only if PD-L1 results are represented as a numeric combined positive score

PD-L1 Results Presented as Summary Assessment

[pdl1_type___4], [pdl1_type_2___4], [pdl1_type_3___4]

Value (Character)

- Summary Assessment

Description

- Indicates that PD-L1 test results are represented as a summary assessment
- Populated only if PD-L1 results are represented as a summary assessment

Any Positive PD-L1 Result Reported on Pathology Report

[pdl1_positive_any]

Value (Character)

- Yes
- No

Description

- Indicates whether any PD-L1 testing on a given pathology report returned a positive result
- Based on any of the following criteria being met:
  - A low, high, or positive summary score
  - A percentage or percentage range tumor cells greater than 0
  - A percentage or percentage range of infiltrating immune cells greater than 0
- Populated only if PD-L1 testing is indicated on pathology report (i.e. [pdl1_testing] = "Yes")

Percentage of Tumor Cells Positive for PD-L1

[pdl1_perc], [pdl1_perc_2], [pdl1_perc_3]

Value (Numeric)

- 0-100

Description

- The percentage of tumor cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of tumor cells ([pdl1_type___1], [pdl1_type_2___1], [pdl1_type_3___1] = "Percentage or Percentage Range of Tumor Cells")

Lower Range (%) of Tumor Cells Positive for PD-L1

[pdl1_tclrange], [pdl1_tclrange_2], [pdl1_tclrange_3]

Value (Character)

- <1
- 0-100
- Other
- Not Applicable

Description

- Lower range (%) of tumor cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of tumor cells ([pdl1_type___1], [pdl1_type_2___1], [pdl1_type_3___1] = "Percentage or Percentage Range of Tumor Cells")

Upper Range (%) of Tumor Cells Positive for PD-L1

[pdl1_tcurange], [pdl1_tcurange_2], [pdl1_tcurange_3]

Value (Character)

- <1
- 0-100
- Other
- Not Applicable

Description

- Upper range (%) of tumor cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of tumor cells ([pdl1_type___1], [pdl1_type_2___1], [pdl1_type_3___1] = “Percentage or Percentage Range of Tumor Cells”)

Percentage of Infiltrating Immune Cells Positive for PD-L1

[pdl1_icperc], [pdl1_icperc_2], [pdl1_icperc_3]

Value (Numeric)

- 0-100

Description

- The percentage of infiltrating immune cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of infiltrating immune cells ([pdl1_type___2], [pdl1_type_2___2], [pdl1_type_3___2] = “Percentage or Percentage Range of Infiltrating Immune Cells”)

Lower Range (%) of Infiltrating Immune Cells Positive for PD-L1

[pdl1_iclrange], [pdl1_iclrange_2], [pdl1_iclrange_3]

Value (Character)

- <1
- 0-100
- Other
- Not Applicable

Description

- Lower range (%) of infiltrating immune cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of infiltrating immune cells ([pdl1_type___2], [pdl1_type_2___2], [pdl1_type_3___2] = “Percentage or Percentage Range of Infiltrating Immune Cells”)

Upper Range (%) of Infiltrating Immune Cells Positive for PD-L1

[pdl1_icurange], [pdl1_icurange_2], [pdl1_icurange_3]

Value (Character)

- <1
- 0-100
- Other
- Not Applicable

Description

- Upper range (%) of infiltrating immune cells positive for PD-L1
- Populated only if PD-L1 testing reported as percentage or percentage range of infiltrating immune cells ([pdl1_type___2], [pdl1_type_2___2], [pdl1_type_3___2] = "Percentage or Percentage Range of Infiltrating Immune Cells")

Numeric Combined Positive Score (CPS) for PD-L1

[pdl1_num], [pdl1_num_2], [pdl1_num_3]

Value (Numeric)

Description

- The Combined Positive Score (CPS) value
- The minimum CPS is 0
- Populated only if PD-L1 testing reported as a numeric combined positive score ([pdl1_type___3], [pdl1_type_2___3], [pdl1_type_3___3] = "Numeric (Combined Positive Score)")

Lower Range of the Combined Positive Score (CPS) for PD-L1

[pdl1_lcpsrange], [pdl1_lcpsrange_2], [pdl1_lcpsrange_3]

Value (Character)

- <1
- 0-11
- Other
- Not Applicable

Description

- The lower range of the Combined Positive Score (CPS)
- Populated only if PD-L1 testing reported as a numeric combined positive score ([pdl1_type___3], [pdl1_type_2___3], [pdl1_type_3___3] = "Numeric (Combined Positive Score)")

Upper Range of the Combined Positive Score (CPS) for PD-L1

[pdl1_ucpsrange], [pdl1_ucpsrange_2], [pdl1_ucpsrange_3]

Value (Character)

- <1
- 0-11
- Other
- Not Applicable

Description

- The upper range of the Combined Positive Score (CPS)
- Populated only if PD-L1 testing reported as a numeric combined positive score ([pdl1_type___3], [pdl1_type_2___3], [pdl1_type_3___3] = "Numeric (Combined Positive Score)")

Summary Assessment of PD-L1

[pdl1_sum], [pdl1_ sum _2], [pdl1_ sum _3]

Value (Character)

- High
- Low
- Positive
- Negative
- Indeterminate/Not stated

Description

- Overall summary assessment value of PD-L1 as stated in pathology report; not specific to immune cells or tumor cells
- Populated only if PD-L1 testing reported as a summary score ([pdl1_type___4], [pdl1_type_2___4], [pdl1_type_3___4] = "Summary Assessment")

Release Version Number

[release_version]

Value (Character)

- 2.0-public

Description

- Indicates the version number of the data release

PRISSMM IMAGING DATASET

The PRISSMM Imaging dataset is structured as one record per imaging report, per patient. Imaging reports were curated beginning within 30 days of the first BPC Project cancer diagnosis. All subsequent imaging reports were recorded (including imaging reports corresponding to non-BPC Project Cancers and subsequent BPC Project Cancer diagnoses; Appendix 1).

The PRISSMM Imaging dataset can be linked to all datasets using the variables [cohort] and [record_id].

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The preferred variable is noted in the description.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, “GENIE”; the second component is the institution’s abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Imaging Report Number

[scan_number]

Value (Numeric)

Description

- Unique identifier for imaging reports based on the date of the scan [image_scan_int]
- Different scans occurring on the same date have distinct scan numbers (e.g. a CT and MRI occurring on the same date may be numbered as scan 1 and scan 2, though the ordering is arbitrary)

Time (Days) from Date of Birth to Imaging Date

[image_scan_int]

Value (Numeric)

Description

- Interval in days from date of birth to scan date

Time (Days, Months, Years) from First BPC Project Cancer to Imaging Date

[dx_scan_days], [dx_scan_mos], [dx_scan_yrs]

Value (Numeric)

Description

- Interval in days [dx_scan_days]; months [dx_scan_mos]; or years [dx_scan_yrs] from first BPC Project Cancer diagnosis to scan date

Time (Days) from Date of Birth to Reference Imaging Date

[image_ref_scan_int]

Value (Numeric)

Description

- Interval days in from date of birth to reference imaging date
- Populated only if there is evidence of cancer on the imaging report ([image_ca] = “Yes, the Impression states or implies there is evidence of cancer”)

Time (Days, Months, Years) from First BPC Project Cancer to Reference Imaging Date

[dx_ref_scan_days], [dx_ref_scan_mos], [dx_ref_scan_yrs]

Value (Numeric)

Description

- Interval in days [dx_ref_scan_days]; months [dx_ref_scan_mos]; or years [dx_ref_scan_yrs] from first BPC Project Cancer diagnosis to reference imaging date

Institution Where Scan was Performed

[image_inst_perf]

Value (Character)

- Internal institution
- External institution

Description

- Indicates whether scan was performed at the internal (i.e. the same institution as the [institution] variable) or external institution.

Institution Where Image was Interpreted

[image_inst_inter]

Value (Character)

- Internal institution
- External institution

Description

- Indicates whether image was interpreted at the internal (i.e. the same institution as the [institution] variable) or external institution.
- Populated only if imaging was performed at an external institution ([image_inst_perf] = "External institution")

Imaging Scan Type

[image_scan_type]

Value (Character)

- CT
- MRI
- PET or PET-CT
- Bone Scan
- Other Nuclear Medicine Scan
- Mammogram – Use for Breast Cancer only
- Other CA-specific scan

Description

- Type of imaging scan

Imaging Site: Brain/Head

[image_scansite___1]

Value (Character)

- Brain/Head

Description

- Indicates a scan of the brain/head
- Populated only if scan is of brain/head

Imaging Site: Spine

[image_scansite___2]

Value (Character)

- Spine

Description

- Indicates a scan of the spine
- Populated only if scan is of the spine

Imaging Site: Neck

[image_scansite___3]

Value (Character)

- Neck

Description

- Indicates a scan of the neck
- Populated only if scan is of the neck

Imaging Site: Chest

[image_scansite___4]

Value (Character)

- Chest

Description

- Indicates a scan of the chest
- Populated only if scan is of the chest

Imaging Site: Abdomen

[image_scansite___5]

Value (Character)

- Abdomen

Description

- Indicates a scan of the abdomen
- Populated only if scan is of the abdomen

Imaging Site: Pelvis

[image_scansite___6]

Value (Character)

- Pelvis

Description

- Indicates a scan of the pelvis
- Populated only if scan is of the pelvis

Imaging Site: Extremity

[image_scansite___7]

Value (Character)

- Extremity

Description

- Indicates a scan of an extremity
- Populated only if scan is of an extremity

Imaging Site: Full body

[image_scansite___8]

Value (Character)

- Full body

Description

- Indicates a full body scan
- Populated only if a full body scan

Imaging Sites

[scan_sites]

Value (Character)

Description

- List of sites scanned on this imaging report
- Concatenation of sites scanned ([image_scansite___1] – [image_scansite___8])

Radiologist Assessment of any Evidence of Cancer on this Imaging Report

[image_ca]

Value (Character)

- Yes, the Impression states or implies there is evidence of cancer
- No, the Impression states or implies there is no evidence of cancer
- The Impression is uncertain, indeterminate, or equivocal
- The Impression does not mention cancer

Description

- Indicates whether the radiologist assessment indicates any evidence of cancer

Radiologist Assessment of Change in Cancer Status

[image_overall]

Value (Character)

- Improving/Responding
- Stable/No change
- Mixed
- Progressing/Worsening/Enlarging
- Not stated/Indeterminate

Description

- Radiologist's assessment of the change in the patient's cancer status
- Populated only if there is evidence of cancer on the imaging report ([image_ca] = "Yes, the Impression states or implies there is evidence of cancer")

Location of Cancer Based on Imaging Report

[image_casite1]-[image_casite15]

Value (Character)

- ICD-O-3 topography code

Description

- Location of cancer on imaging report based on ICD-O-3 topography code
- Populated only if there is evidence of cancer on the imaging report ([image_ca] = "Yes, the Impression states or implies there is evidence of cancer")

Release Version Number

[release_version]

Value (Character)

- 2.0-public

Description

- Indicates the version number of the data release

PRISSMM MEDICAL ONCOLOGIST ASSESSMENT DATASET

The PRISSMM Medical Oncologist Assessment dataset is structured as one record per curated medical oncologist assessment, per patient. Medical oncologist assessments were curated beginning with the month and year of the first BPC Project Cancer diagnosis (Appendix 1). One medical oncologist assessment per month was curated; curation instructions regarding the selection of the assessment to curate are provided in Appendix 3.

The PRISSMM Medical Oncologist Assessment dataset can be linked to all datasets using the variables [cohort] and [record_id].

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The preferred variable is noted in the description.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, “GENIE”; the second component is the institution’s abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Medical Oncologist Visit Identifier

[md_visit_number]

Value (Numeric)

Description

- Unique identifier for curated medical oncologist assessments based on the visit date [md_onc_visit_int]
- Only one medical oncologist assessment per month was curated (Appendix 3)

Time (Days) from Date of Birth to Medical Oncologist Visit

[md_onc_visit_int]

Value (Numeric)

Description

- Interval in days from date of birth to date of medical oncologist visit
- Based on the date the visit occurred, not the date the assessment was signed or uploaded.

Time (Days, Months, Years) from First BPC Project Cancer to Medical Oncologist Visit

[dx_md_visit_days], [dx_md_visit_mos], [dx_md_visit_yrs]

Value (Numeric)

Description

- Interval in days [dx_md_visit_days]; months [dx_md_visit_mos]; or years [dx_md_visit_yrs] from first BPC Project Cancer diagnosis to medical oncologist visit.
    - Based on the date the visit occurred, not the date the assessment was signed or uploaded.

Cancer Diagnosis Assessed by Medical Oncologist

[md_type_ca_cur]

Value (Character)

- Adrenocortical Carcinoma
- Anal Cancer
- Appendix Cancer
- Bile Duct Cancer
- Bladder Cancer
- Brain Cancer
- Breast Cancer
- NET or Carcinoid
- Cervical Cancer
- Colon Cancer
- Colon/Rectum Cancer
- Esophagus Cancer
- Ewing Sarcoma
- Fallopian Tube Cancer
- Gallbladder Cancer
- Germ Cell Tumor
- GIST
- Head and Neck Cancer
- Mesothelioma
- Ill Defined/Cancer of Unknown Primary
- Liver Cancer
- Lung Cancer, NOS
- Melanoma
- Merkel Cell
- Neuroblastoma
- Non Small Cell Lung Cancer
- Osteosarcoma
- Ovarian Cancer
- Pancreatic Cancer
- Parathyroid Cancer
- Penis Cancer
- Peritoneum Cancer
- Placenta Cancer
- Prostate Cancer
- Rectum and Rectosigmoid Cancer
- Renal Kidney Cancer
- Renal Pelvis Cancer
- Retinoblastoma
- Rhabdomyosarcoma
- Scrotum Cancer
- Small Cell Lung Cancer
- Small Intestine Cancer
- Stomach Cancer
- Testis Cancer
- Thymus Cancer
- Thyroid Cancer
- Uterus Cancer
- Vagina Cancer
- Vulva Cancer
- Wilms Tumor
- Other

Description

- Indicates the cancer diagnosis associated with the given medical oncologist assessment

Medical Oncologist Assessment of Evidence of Cancer

[md_ca]

Value (Character)

- Yes, the Impression/Plan states or implies there is evidence of cancer
- No, the Impression/Plan states or implies there is no evidence of cancer
- Impression/Plan is uncertain, indeterminate, or equivocal
- Impression/Plan does not mention cancer

Description

- Medical oncologist's assessment of whether there is evidence of cancer

Medical Oncologist Assessment of Change in Cancer Status

[md_ca_status]

Value (Character)

- Improving/Responding
- Stable/No change
- Mixed
- Progressing/Worsening/Enlarging
- Not stated/Indeterminate

Description

- Medical oncologist's assessment of the change in the patient's cancer status
- Populated only if there is evidence of cancer on the medical oncologist assessment ([md_ca] = "Yes, the Impression/Plan states or implies there is evidence of cancer")

Release Version Number

[release_version]

Value (Character)

- 2.0-public

Description

- Indicates the version number of the data release

CANCER PANEL TEST (NEXT GENERATION SEQUENCING) DATASET

The Cancer Panel Test refers to high-throughput next generation sequencing (NGS) that has been performed through multi-gene panels. The Cancer Panel Test dataset is structured as one record per cancer panel test and associated cancer diagnosis, per patient. For example, if a cancer panel test was definitively associated with one cancer diagnosis, there will be one corresponding record in this dataset. If the curator was unsure of which diagnosis out of two cancer diagnoses corresponded to the cancer panel test, there will be two corresponding records in this dataset: one for the first potentially associated cancer diagnosis and another for the second potentially associated cancer diagnosis. All BPC Project Cancers have an associated cancer panel test (Appendix 1).

The terms "cancer panel test (CPT)" and "next generation sequencing (NGS)" are used interchangeably.

The Cancer Panel Test dataset can be linked to the following datasets:

- *BPC Project Cancer Diagnosis and Cancer-Directed Regimen datasets using the variables [cohort], [record_id] and [ca_seq]*
- *PRISSMM Pathology dataset using [cohort], [record_id], [path_proc_number] and [path_report_number]*
- *Patient Characteristics, PRISSMM Imaging, and PRISSMM Medical Oncologist Assessment datasets using [cohort], [record_id]*
- *Cannot be linked to the Non-BPC Project Cancer Diagnosis dataset because non-BPC Project Cancer diagnoses were not genomically sequenced (Appendix 1)*

Field names shaded in gray indicate that an alternative variable is preferred for analysis. The preferred variable is noted in the description.

BPC Project Cohort

[cohort]

Value (Character)

- NSCLC

Description

- Indicates the BPC Project Cancer type
- Primary key for merging across all datasets

Record ID

[record_id]

Value (Character)

- GENIE-[INSTITUTION]-XXXXXX

Description

- De-identified, unique patient ID
- Conforms to the following the convention: GENIE-[INSTITUTION]-XXXX. The first component is the string, "GENIE"; the second component is the institution's abbreviation; the third component is a unique ID for the patient.
- Primary key for the AACR Project GENIE genomic datasets [PATIENT_ID].

Institution

[institution]

Value (Character)

- DFCI = Dana Farber Cancer Institute
- MSK = Memorial Sloan Kettering Cancer Center
- VICC = Vanderbilt Ingram Cancer Center
- UHN = University Health Network (Princess Margaret)

Description

- Indicates the patient's internal institution of cancer care
- Corresponds to variable [center] in AACR Project GENIE data

Cancer Panel Test (Next Generation Sequencing) Number

[cpt_number]

Value (Numeric)

- 1 = First curated next generation sequencing (NGS) test for this patient
- 2 = Second curated NGS test for this patient
- … n = nth NGS test for this patient

Description

- Order for the curated next generation sequencing (NGS) test based on the report date

Cancer Diagnosis Number Associated with Next Generation Sequencing Test

[ca_seq]

Value (Numeric)

- 0 = first and only cancer
- 1 = first of two or more primaries
- 2 = second of two or more primaries
- …10 = tenth of ten or more primaries

Description

- The cancer sequence number associated with this NGS
- If the NGS test cannot be definitively associated with a single cancer diagnosis, this variable represents the cancer sequence of each potentially associated cancer diagnosis. If that is the case, each potentially associated cancer diagnosis will be a separate row in the dataset.
- Populated only if the diagnosis associated with the NGS test is known
- Primary key for the Cancer Diagnosis, Cancer-Directed Regimen and Cancer Panel Test datasets

Number of Cancer Diagnoses Associated with Next Generation Sequencing Test

[cpt_n_ca_seq]

Value (Numeric)

Description

- The number of associated cancer diagnoses for this NGS test
- If >1 then the NGS test cannot be definitively associated with a single cancer diagnosis

Time (Days) from Date of Birth to Cancer Panel Test Order Date

[cpt_order_int]

Value (Numeric)

Description

- Interval in days from date of birth to the date that the cancer panel test was ordered
- This variable is not available for all institutions

Year of Next Generation Sequencing

[cpt_seq_date]

Value (Character)

Description

- Year of NGS

Derived Time (Days, Months, Years) from Date of Birth to Sequencing Report

[dob_cpt_report_days], [dob_cpt_report_mos], [dob_cpt_report_yrs]

Value (Numeric)

Description

- Interval in days [dob_cpt_report_days], months [dob_cpt_report_mos] and years [dob_cpt_report_yrs] from date of birth to date of sequencing report

Next Generation Sequencing Test Report Returned On or After Date of Death

[cpt_report_post_death]

Value (Numeric)

- 1 = Yes, cancer panel test was returned on or after patient's date of death
- 0 = No, cancer panel test was not returned on or after patient's date of death

Description

- Indicates whether cancer panel test report was returned on or after patient's date of death

Next Generation Sequencing Test Report Returned On or After Date Patient Last Known Alive

[cpt_report_post_last_alive]

Value (Numeric)

- 1 = Yes, cancer panel test was returned on or after patient's last known alive date
- 0 = No, cancer panel test was not returned on or after patient's last known alive date

Description

- Indicates whether cancer panel test report was returned on or after patient's last known alive date
- Populated only if patient was not known to be dead at the time of curation. For an indicator for whether the cancer panel test report was returned after death, see variable [cpt_report_post_death].

Time (Days, Months, Years) from Diagnosis to Sequencing Report

[dx_cpt_rep_days], [dx_cpt_rep_mos], [dx_cpt_rep_yrs]

Value (Numeric)

Description

- Interval in days [dx_cpt_rep_days]; months [dx_cpt_rep_mos]; or years [dx_cpt_rep_yrs] from cancer diagnosis to sequencing report

Time (Days, Months, Years) from Diagnosis to Pathology Procedure Corresponding to the Next Generation Sequencing Report

[dx_path_proc_cpt_days], [dx_path_proc_cpt_mos], [dx_path_proc_cpt_yrs]

Value (Numeric)

Description

- Interval in days [dx_path_proc_cpt_days]; months [dx_path_proc_cpt_mos]; or years [dx_path_proc_cpt_yrs] from cancer diagnosis to date of pathology procedure corresponding to cancer panel test

Time (Days, Months, Years) from Pathology Procedure to the Next Generation Sequencing Report Date

[path_proc_cpt_rep_days], [path_proc_cpt_rep_mos], [path_proc_cpt_rep_yrs]

Value (Numeric)

Description

- Interval in days [path_proc_cpt_rep_days]; months [path_proc_cpt_rep_mos]; or years [path_proc_cpt_rep_yrs] from date of pathology procedure corresponding to cancer panel test to date of cancer panel test report

Pathology Procedure Number of Next Generation Sequencing Specimen

[path_proc_number]

Value (Numeric)

Description

- Pathology procedure in which the specimen is described
- Primary key for PRISSMM Pathology and Cancer Panel Test datasets

Pathology Report Number of Next Generation Sequencing Specimen

[path_rep_number]

Value (Numeric)

Description

- Pathology report in which the specimen is described
- Primary key for PRISSMM Pathology and Cancer Panel Test datasets

GENIE Sample ID

[cpt_genie_sample_id]

Value (Character)

Description

- GENIE sample ID corresponding to specimen
- Corresponds to variable [sample_id] in AACR Tier 1 data

Next Generation Sequencing Specimen OncoTree Diagnosis Code

[cpt_oncotree_code]

Value (Character)

Description

- The primary cancer diagnosis code based on the OncoTree ontology
- Corresponds to variable [oncotree_code] in AACR Tier 1 data

Data Standard: OncoTree Ontology

Specimen Sample Type

[sample_type]

Value (Character)

- Primary tumor
- Lymph node metastasis
- Distant organ metastasis
- Metastasis site unspecified
- Local recurrence
- Not otherwise specified
- Not applicable or hematologic malignancy

Description

- Sample type associated with specimen on which NGS was performed
- Corresponds to variable [sample_type] in AACR Tier 1 data

Sequencing Assay ID

[cpt_seq_assay_id]

Value (Character)

Description

- The institutional assay identifier for the NGS genomic testing platform.
- Components are separated by hyphens, with the first component corresponding to the institution's abbreviation.
- All specimens tested by the same platform should have the same identifier.
- Corresponds to variable [seq_assay_id] in AACR Tier 1 data

Release Version Number

[release_version]

Value (Character)

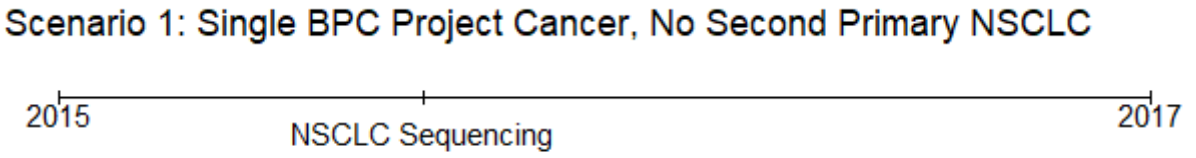
- 2.0-public

Description

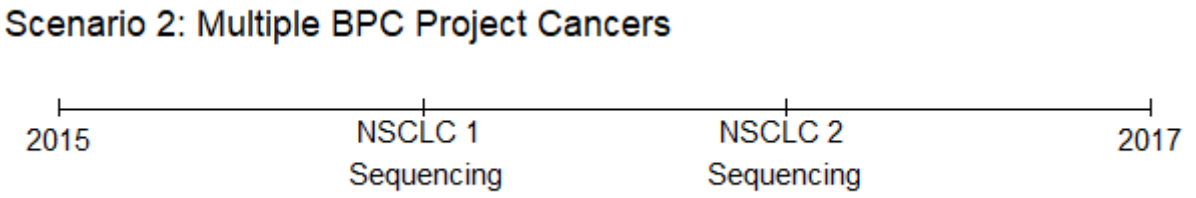
- Indicates the version number of the data release

Appendix 1. BPC Project and Non-BPC Project Cancers

Definition of BPC Project Cancer: A BPC Project Cancer is the cancer that met the eligibility criteria for the project (i.e. genomic sequencing reported). The terms "BPC Project Cancer" and "index cancer" are used interchangeably. In Scenario 1, the patient had a single eligible cancer with associated genomic sequencing.



Some patients may have more than one BPC Project Cancer because they have multiple sequenced cancers that met the eligibility criteria. For example, in Scenario 2, a patient with a diagnosis of NSCLC that was sequenced in 2015 and a second primary of NSCLC that was sequenced in 2017 will have each diagnosis classified as a BPC Project Cancer in the NSCLC BPC Project.

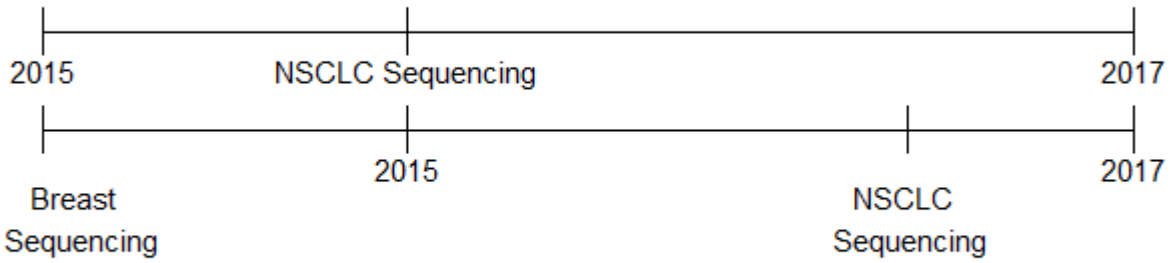


Definition of Non-BPC Project Cancer: A non-BPC Project Cancer diagnosis can be a diagnosis of the same or different cancer type as the BPC Project Cancer that occurs prior to, simultaneous with, or after the BPC Project Cancer. Non-BPC Project Cancers curated included: 1) non-BPC Project invasive cancer and in situ/non-invasive cancer diagnoses, and 2) other benign tumors. In a very small number of instances, some benign tumors (behavior codes 0 and 1), such as a benign brain tumor or a hemangioma, are obtained from the institution's tumor registry. These tumors were ineligible for curation, in which case only the tumor registry data is available.

The non-BPC Project Cancers do not have associated genomic sequencing.

A non-BPC Project cancer diagnosis could be a different cancer type altogether, such as a breast cancer diagnosis in a patient included in the NSCLC BPC Project. The NSCLC cancer is the cancer that made the patient eligible for the BPC Project, but some information regarding the breast cancer diagnosis was curated. In Scenario 3, the breast cancer that was sequenced is classified as a non-BPC Project cancer for the NSCLC BPC Project.

Scenario 3: Single BPC Project Cancer, Single Non-BPC Project Cancer



Appendix 2. Cancer-Directed Drugs

- Mercaptopurine(AltiMercaptopurine,Azathiopurine,BW57323H,Flocofil,Ismipur,Leukerin,Leupurin,Mern,NCIC04886,PuriNethol,Purimethol,Purinethiol,Purinethol,U4748,WR2785)
- Acetylcysteine(Acetadote,Airbron,Broncholysin,Brunac,Fabrol,Fluatox,Fluimucetin,Fluimucil,Fluprowit,Muco Sanigen,Mucocedyl,Mucolator,Mucolyticum,Mucomyst,Mucosolvin,Mucret,NAC,NeoFluimucil,Parvolex,Respaire,Tixair)
- Recombinant Interferon Alfa(Alferon,Alpha Interferon,Leukocyte Interferon,Lymphoblast Interferon,Lymphoblastoid Interferon)
- Aminoglutethimide(Cytadren,Aminoblastin,Elipten,Orimenten,Rodazol)
- Asparaginase(ASP1,Asparaginase II,Asparaginase E.Coli,Colaspase,Elspar,Kidrolase,LASP,LAsnase,Laspar,Leucogen,Leunase,MK965,Paronal,Re82TAD15,Serasa,Spectrila)
- Azacitidine(5 AZC,5AC,Mylosar,U18496,Vidaza)
- BCG Vaccine(BCG TICE,Bacille CalmetteGuerin Live,ImmuCyst,Imovax BCG,Monovax,Oncotice,Pacis,Pastimmun)
- Bleomycin(BLEO,BLM)
- Busulfan(BUS,Busulfanum,Busulfex,CB2041,GT41,Glyzophrol,Joacamine ester,Mielucin,Misulban, Misulfan,Mitosan,Myeleukon,Myelosan, Mylecytan,Myleran,Sulfabutin,WR19508)
- Captopril,Capoten
- Carmustine(BCNU,Becenum,BiCNU, Carmubris,Carmustin,Carmustinum,FDA 0345,Nitrourean,Nitrumon,K27702,SRI1720,WR139021)
- Chlorambucil,Chlorambucilum,Chloraminophen,Chloraminophene,Chlorbutine,Chlorobutin,ChlorobutineEcloril,Leukeran,Leukersan,Leukoran,Linfolizin,Linfolysin
- Cimetidine(Tagamet)
- Cisplatin(Cisplatina,Cisplatinum,Cisplatyl, Citoplatino,Citosin,Cysplatyna,DDP,Lederplatin,Metaplatin,Neoplatin,Peyrone's Chloride,Peyrone's Salt,Placis,Plastistil,Platamine,Platiblastin,PlatiblastinS,Platinex,Platinol,Platinoxan, Platinum Diamminodichloride,Platiran,Platistin,Platosin)
- Clomiphene Citrate(Serophene)
- Cyclophosphamide(CTX,CYCLOcell,Carloxan,Clafen,Claphene,Cyclophosphamide,Cyclophosphan,Cyclophosphane,Cyclostine,Cytophosphan,Cytophosphane,Cytoxan,Endoxan,Fosfaseron,Genoxal,Genuxal,Ledoxina,Mitoxan,Neosar,Revimmune,Syklofosfamid,WR138719)
- Cytarabine(ARAcell,Alexan,AraC,Arabine, Aracytidine,Aracytin,Aracytine,CHX3311, Cytarabinum,Cytarbel,Cytosar,CytosarU, Cytosine Arabinoside,Erpalfa,Starasid,Tarabine PFS,U 19920,U19920,Udicil,WR28453)
- Dacarbazine(Asercit,Biocarbazine,DIC,DTIC,DTICDome,Dacarbazina,Dacarbazine DTIC,Dacatic,Dakarbazin,Deticene,Detimedac,Fauldetic,Imidazole Carboxamide,WR139007
- Dactinomycin(Actinomycin A IV,Actinomycin C1,Actinomycin D,Actinomycin I1,Actinomycin IV,Actinomycin X 1,Cosmegen,DACT,Lyovac Cosmegen)
- Epirubicin Hydrochloride(Ellence,IMI28,Pharmorubicin PFS)
- Estramustine(Leo 275,RO 218837)
- Amifostine(APAETP,Cytofos,Ethiofos,Ethyol,Gammaphos,WR2721,YM08310)
- Etoposide(EPEG,Lastet,Toposar,VP 16213,VP16,VP16213,Vepesid)
- Floxuridine(FDUR,FUDR,WR138720)
- Fluorouracil(5FU,5Fluorouracil,5Fluracil,AccuSite,Adrucil,Carac,Fluouracil,Flurablastin,Fluracedyl,Fluracil,Fluril,Fluroblastin,Ribofluor,Ro29757)
- Flutamide(Apimid,Cebatrol,Chimax,Cytomid,Drogenil,Euflex,Eulexin,Eulexine,FLUT,Flucinom,Flucinome,Flugerel,Fluken,Flulem,FlutaGry,Flutabene,Flutacan,Flutame

x,Flutamin,Flutan,Flutaplex,Fugerel,Grisetin,Niftolide,Oncosal,Profamid,Prostacur,Prostadirex,Prostica,Prostogenat,SCH 13521,Tafenil,Tecnoflut,Testotard)

- Altretamine(ENT50852,HMM,HXM,Hemel,Hexalen,Hexaloids,Hexastat)
- Hydroxyurea(Droxia,Hydrea,Hydroxycarbamide,Litalir,OncoCarbideOxeron,SQ1089,Syrea,WR83799)
- Ifosfamide(Cyfos,Holoxan,Holoxane,IFO,IFOCell,IFX,Ifex,Ifolem,Ifomida,Ifomide,Ifosfamidum,Ifoxan,MJF9325,Mitoxana,Naxamide,Seromida,Tronoxal,Z4942)
- Isotretinoin(Absorica,Accure,Accutane,Amnesteem,Cistane,Claravis,Isotrex,Isotrexin,Myorisan,Neovitamin A,Oratane,Ro43780,Roaccutan,Roaccutane,Roacutan,Sotret,ZENATANE)
- Leucovorin Calcium(Calfolex,Calinat,Cehafolin,Citofolin,Citrec,Citrovorum Factor,Cromatonbic Folinico,Dalisol,Disintox,Divical,Ecofol,Emovis,FOLIcell,Flynoken A,Folaren,Folaxin,Foliben,Folidan,Folidar,Folinac,Folinate Calcium,Folinoral,Folinvit,Foliplus,Folix,Imo,Lederfolat,Lederfolin,Leucosar,Rescufolin,Rescuvolin,Tonofolin,Wellcovorin)
- Lomustine(Belustin,CCNU,Cecenu,CeeNU,Citostal,Gleostine,Lomeblastin,Lomustinum,Lucostin,Lucostine,Prava,RB1509,WR 139017)
- Mechlorethamine Hydrochloride(Caryolysine,Chlorethamine HCl,Chlorethazine Hydrochloride,Chloromethine HCl,Cloramin,Erasol,HN 2 Hydrochloride,Mustargen,Mustargen HCl,Mustargen Hydrochloride,Mustine Hydrochloride,NLost,OncoCloramin,WR147650)
- Melphalan(L Sarcolysine,Melphalanum,Phenylalanine Mustard,Sarcoclorin,Sarkolysin,WR19813)
- Methotrexate(Abitrexate,AlphaMethopterin,Amethopterin,Brimexate,CL14377,Emtexate,Farmitrexat,Fauldexato,Folex,Lantarel,Ledertrexate,Lumexon,MTX,Maxtrex,Medsatrexate,Metex,Methoblastin, Methylaminopterin,Methotrexatum,Metotrexato,Metrotex,Mexate,Novatrex,Texate,Tremetex,Trexeron,Trixilem,WR19039)
- Plicamycin(A2371,Aureolic acid,MTH,Mithracin,Mithracine,Mithramycin,PA144)
- Mitotane(CB313,Chloditan,Chlodithane,DDD,Mytotan,WR1304)5
- Mitoxantrone Hydrochloride(CL232315,DHAD,DHAQ,DHAQ,Mitroxone,Neotalem,Novantrone,Onkotrone,Pralifan)
- Pentostatin(CI825,CoVidarabine,Covidarabine,DCF,Deoxycoformycin,Nipent,PD81565)
- Procarbazine Hydrochloride(MIH hydrochloride,Matulane,NCIC01810,Natulan,Natulanar,PCB,PCZ,Ro46467 1)
- Streptozocin(STZ,U 9889,U9889,Zanosar)
- Sulindac(Aflodac,Algocetil,ApoSulin,Arthrocine,Artribid,Citireuma,Clinoril,Clisundac,Imbaral,MK231,NovoSundac,Reumofil, Reumyl,Sulinol,Sulreuma)
- Tamoxifen Citrate(ApoTamox,Clonoxifen,Dignotamoxi,Ebefen,Emblon,Estroxyn,Fentamox,GenTamoxifen,Genox,ICI 46,474,ICI46474,Jenoxifen,Kessar,Ledertam,Lesporene,Nolgen,Noltam,Nolvadex, NolvadexD,Nourytam,NovoTamoxifen,Novofen,Noxitem,Oestrifen,Oncotam,PMSTamoxifen,Soltamox,TAM,Tamax,Tamaxin, Tamifen,Tamizam,Tamofen,Tamoxasta,Tamoxifeni Citras,Zemide)
- Teniposide(EPT,PTG,Thenylidene Lignan,VM 26,VM26,Vehem,Vumon)
- Thalidomide(Contergan,Distaval,Kevadon,Neurosedyn,Pantosediv,Pantosediv,Sedalis,Sedoval K17,Softenon,Softenon,Synovir,Talimol,Thalomid
- Thiotepa(Oncotiotepa,STEPA,TESPA,TIOTEF,TSPA)

- Thioguanine(2Amino 6MP,BW 5071,Lanvis,Tabloid,Tioguanin,WR1141,U3B,X27)
- Tretinoin(ATRA,Aberel,Airol,Airol,Aknoten,Avita,Cordes Vas,Dermairol,EpiAberel,Eudyna,Renova,RetinA,RetinA MICRO,RetisolA,Ro 5488,StievaA,StievaA Forte,Trans Retinoic Acid)
- Trifluridine(F3TdR,Triflorothymidine)
- Coenzyme Q10(CoQ10,CoQ10,Coenzyme Q10,UBIDECARENONE,Ubiquinone 10,coenzyme Q10,vitamin Q10)
- Verapamil
- Vinblastine Sulfate,29060LE,Exal,VINCALEUKOBLASTINE,Velban,Velbe,Velsar
- Gemcitabine Hydrochloride(FF 10832,FF10832,FF10832,Gemzar,LY188011,LY188011,dFdCyd)
- SarCNU(sarcosinamide nitrosourea)
- 2Methoxyestradiol(2MeE2,2MeOE2,2Methoxy Estradiol,2ME2,Panzem)
- Decitabine(Dacogen,Deoxyazacytidine,Dezocitidine)
- Acitretin(Etretin,Neotigason,Ro 101670,Soriatane
- Arsenic Trioxide(ATO,Trisenox)
- Bryostatin 1(B705008K112)
- Clarithromycin(Abbott56268,Biaxin)
- Etoposide Phosphate(Etopophos)
- Exemestane(Aromasin,FCE24304)
- Fenretinide(HPR4)
- Fludarabine Phosphate(Beneflur,Fludara,Oforta,SH T 586)
- Omacetaxine Mepesuccinate(CGX635,Ceflatonin,HHT,Synribo)
- Mafosfamide
- Megestrol Acetate(Maygace,Megace,Megestat,Megestil,Niagestin,Ovaban,Pallace,SC10363)
- Nilutamide(Anandron,Nilandron,RU23908)
- Oxaliplatin(1OHP,Ai Heng,DACPLAT,Dacotin,ELOXATIN,Eloxatine,JM83)
- Pegaspargase(Oncaspar,PEGLA)
- Temozolomide(M & B 39831,Methazolastone,RP46161,SCH52365,Temcad,Temodal,Temodar,Temomedac)
- Toremifene(Farestone)
- Triptorelin(AY25650,CL118532,Decapeptyl,Detryptoreline)
- Carboplatin(Blastocarb,CBDCA,Carboplat,Carboplatin Hexal,Carboplatino,Carboplatinum,Carbosin,Carbosol,Carbotec,Displata,Ercar,JM8,Nealorin,Novoplatinum,Paraplat,Paraplatin,Paraplatine,Platinwas,Ribocarbo)
- AGM1470,TNP470
- O6Benzylguanine
- Leuprolide Acetate(A43818,Carcinil,DepoEligard,Enanton,Enantone,EnantoneGyn,Ginecrin,LEUP,Leuplin,Lucrin Depot,Lupron,Lupron Depot,Procren,Procrin,Prostap,TAP144,Trenantone,UnoEnantone,Viadur)
- Doxorubicin Hydrochloride,ADM,Adriacin,Adriamycin,Adriamycin Hydrochloride,Adriamycin RDF,Adriamycine,Adriblastina,Adriblastine,Adrimedac,Chloridrato de Doxorrubicina,DOX,DOXOCELL,Doxorubin,FI 106,FI106,Rubex
- Cladribine,CdA(Leustat,Leustatin,Leustatine,RWJ26251)
- Valrubicin,AD32,AD 32,AD32,Valstar,Valtaxin)
- Mivobulin Isethionate(CI980)
- Fulvestrant(Faslodex,ICI182780,ZD9238)
- Recombinant Interleukin12(Cytotoxic Lymphocyte Maturation Factor,IL12,Interleukin 12,NMIL12,Ro247472)
- Irinotecan Hydrochloride(CPT11,Campto,Camptosar,U101440E)
- Vinorelbine Tartrate(Biovelbin,Eunades,KW2307,NVB,Navelbine,Vinorelbine Ditartrate)
- Paclitaxel(Anzatax,Asotax,Bristaxol,Praxel,Taxol)
- Goserlin Acetate(Zoladex,ZDX)

- Imiquimod(Aldara,R837,S26308,Zyclara)
- Recombinant CD40Ligand,CD154 antigen,CD40L,TBAM,rhu CD40L
- Denileukin Diftitox(DAB389IL2,LY335348,Ontak)
- Rubitecan(Camptogen,Nitrocamptothecin,Orathecin,RFS 2000
- Aldesleukin(Proleukin,Recombinant Human IL2,rserHuIL2)
- Docetaxel (Taxotere,RP56976)
- Letrozole(CGS 20267,Femara)
- Pemetrexed Disodium(Alimta,LY231514)
- Tezacitabine(FMdC,MDL 101,731)
- Romidepsin(Depsipeptide,FK228,FR901228,Istodax)
- Edrecolomab(MOAB171A,Panorex)
- Pegylated Liposomal Doxorubicin Hydrochloride(ATI0918,Caelyx,DOXSL,DOXIL,Doxorubicin HCl Liposomal,Evacet,LipoDox,LipoDox,Lipodox 50,Liposomal Adriamycin,Pegylated Liposomal Doxorubicin Hydrochloride,SLiposomal Doxorubicin,Stealth Liposomal Doxorubicin,TLC D99)
- Dinutuximab(Ch14.18,Unituxin)
- Alvocidib Hydrochloride(Flavopiridol Hydrochloride,HL275,L868275)
- Alitretinoin(ALRT1057,LGD1057,Panretin, Panretyn,Panrexin)
- Eflornithine Hydrochloride,MDL 71782,Ornidyl,RMI71782)
- Daunorubicin Hydrochloride(CERUBIDINE,Ondena,RP13057,Rubidomycin Hydrochloride,Rubilem)
- Idarubicin Hydrochloride(IMI30,Idamycin,SC33428,Zavedos)
- Bicalutamide(Casodex,Cosudex,ICI176334)
- Anastrozole(Arimidex,ICID1033,ZD1033)
- DX521
- KRN5500
- Elinafide(LU79553)
- Becatecarin(BMS181176,BMY2755714,DEAE-Rebeccamycin,NSC655649,Rebeccamycin Analogue,XL119)
- Bexarotene(LGD1069,Targretin)
- Trastuzumab(ABP980,ALT02,Antip185HER2,HER2 Monoclonal Antibody,Herceptin,Herceptin PF05280014,Herzuma, Ogivri, Ontruzant, PF05280014, RO0452317,Trastuzumab ABP 980,Trastuzumab ALT02,Trastuzumab HLX02, Trastuzumab PF05280014, Trastuzumab dkst, cerb2 Monoclonal Antibody, rhuMAb HER2,trastuzumab EG12014, Trastuzumab-pkrb, Trazimera)
- Alemtuzumab(Campath,LDP03,Lemtrada)
- Imatinib Mesylate(CGP57148B, Gleevec, STI571)
- Plitidepsin(APLD, Aplidin,Aplidine,DDB,Dehydrodemnin B)
- Trabectedin(ET743,Ecteinascidin,Yondelis)
- Rituximab(ABP798,BI 695500,C2B8 Monoclonal Antibody,CTP10,IDEC102,IDECC2B8,MabThera,PF05280586,RTXM83,Rituxan,Rituximab ABP 798,Rituximab BI 695500,Rituximab CTP10,Rituximab GB241,Rituximab IBI301,Rituximab PF05280586,Rituximab RTXM83,Rituximab SAIT101,Truxima,rituximab abbs)
- Tipifarnib(R115777,Zarnestra)
- Nelarabine(506U78,Arranon,Compound 506U78,GW506U78)
- Irofulven,HMAF,Hydroxymethylacylfulvene,MGI 114,MGI114
- Cetuximab,Cetuximab CDP1,Cetuximab CMAB009,Cetuximab KL 140,Chimeric MoAb C225,Erbitux,IMCC225
- Vincristine Sulfate,Kyocristine,Leurocristine Sulfate,Leurocristine sulfate,Oncovin,Vincasar,Vincosid,Vincrex)
- Interferon Alfacon1(Advaferon,CIFN,Consensus

Interferon,IFN Alfacon1,Infergen,rmetHuIFNCon1)
- ISIS3521(Affinitac, Affinitak, CGP 64128A, LY900003)
- Capecitabine(Ro091978 000,Xeloda)
- Paclitaxel Poliglumex(CT2103, CT2103, PGTXL, Xyotax)
- Vorinostat(SAHA,Suberanilohydroxamic Acid, Zolinza)
- Gemtuzumab Ozogamicin(CDP771,CMA676,Mylotarg)
- Yttrium Y90 Ibritumomab Tiuxetan(IDECY2B8,Y 90 Zevalin)
- Mitomycin(Ametycine,MITO,MITOC,MITOMYCIN C, MitoMedac, Mitocin, MitocinC,Mitolem,MitomycinX,Mitomycine C,Mitosol,Mitozytrex,Mutamycin,Mutamycine,NCIC04706)
- Lonafarnib(SCH66336,Sarasar,lonafarnib)
- Semaxanib(Semoxind, Sugen 541)
- Tegafurgimeraciloteracil Potassium(BMS247616,S1,STS1,Teysuno)
- Cilengitide(EMD121974)
- L778123
- Temsirolimus(CCI779, Torisel)
- Peginterferon Alfa2b(PEG IFN Alfa2b,PEGIntron,CH54031,Sylatron)
- Bortezomib(LDP341,MLN341,PS341,Velcade)
- Gefitinib(Iressa,ZD1839)
- BMS214662
- Panitumumab(ABXEGF, Vectibix)
- Ribozyme RPI.4610(Angiozyme,AntiFlt1 Ribozyme,RPI.4610,RPI.4610,RPI4610,Ribozyme RPI.4610)
- Vatalanib(CGP79787,PTK787,ZK232934)
- Midostaurin(CGP41251, PKC412,Rydapt)
- Lurtotecan Liposome(NX211,OSI211)
- Canertinib Dihydrochloride(CI1033,PD0183805002B, PD183805)
- Epratuzumab(LymphoCide,hLL2)
- NM3
- Recombinant Interferon Alfa2a(Alpha 2 Interferon,IFN alpha2A,Laroferon,A,rHuIFNa 2a)
- Sipuleucel T(APC8015,PA2024 PAP GMCSFLoaded Dendritic Cell Vaccine, Provenge,SipT)
- Brentuximab(CAC10,SGN30)
- Matuzumab(EMD72000)
- Abarelix(PPI149,Plenaxis,R3827)
- Bevacizumab(AntiVEGF rhuMAb, BEVZ92,Bevacizumab-awwb,Bevacizumab BI 695502,Bevacizumab CBT 124,Bevacizumab FKB238,Bevacizumab HD204,Bevacizumab HLX04,Bevacizumab MIL60,Bevacizumab QL 1101,HD204,Avastin,Mvasi)
- Sorafenib Tosylate(BAY439006 Tosylate,BAY549085,Nexavar)
- Liposomal Daunorubicin Citrate(DaunoXome)
- Interleukin12 Gene(IL12 gene,Pralatrexate,FOLOTYN,PDX)
- Testolactone(Fludestrin,SQ9538,Teslac)
- Etanercept(Enbrel,TNFR:Fc)
- gp100 Antigen(glycoprotein 100)
- Tretinoin Liposome(AR623,All transretinoic acid liposomal,Atragen,Tretinoin Liposomal,TretinoinLF)
- Recombinant Vaccinia PSA Vaccine(rVPSA,rVPSA Vaccine)
- Iodine I 131 Tositumomab(Bexxar)
- Aminocamptothecin Colloidal Dispersion
- Carmustine Implant(BCNU Wafer,Carmustine Copolymer,Carmustine Wafer,Gliadel,Gliadel Wafer)
- Tositumomab(AntiCD20 Antibody,MoAb AntiB1)
- Canfosfamide Hydrochloride(TLK286,Telcyta)
- Methyl 5 Aminolevulinate Hydrochloride Cream(Metvix,Metvixia)
- Ipilimumab(BMS734016,MDX010,MDXCTLA4,Yervoy)
- Edotecarin(J107088)
- Lenalidomide(CC5013,CDC501,Revlimid)
- Recombinant FowlpoxProstate Specific Antigen Vaccine(rFPSA)

- Ziv Aflibercept(AFLIBERCEPT,AVE0005,Eylea, VEGF Trap,Zaltrap)
- Nabpaclitaxel(ABI 007,ABI007,Abraxane,Albumin bound Paclitaxel,Nab paclitaxel,Nanoparticle Albumin bound Paclitaxel)
- Erlotinib Hydrochloride(CP358,774,OSI774,Tarceva)
- Vincristine Sulfate Liposome(Marqibo)
- Recombinant Vaccinia(MUC1 Vaccine,rVMUC1 Vaccine)
- Vandetanib(AZD6474,Caprelsa,ZD6474,ZD6474,Zactima)
- Topotecan Hydrochloride(Hycamptamine, Hycamtin,SKF S104864A)
- ONYX015(CI1042)
- Indisulam(E7070)
- Clofarabine(Clofarex,Clolar)
- Eribulin Mesylate(B1939 Mesylate,E7389,ER086526,Halaven)
- Sunitinib Malate(SU011248, Sutent)
- T900607
- Agatolimod Sodium(CpG 7909,PF3512676,ProMune)
- Lapachone(beta Lapachone)
- MART1 Antigen(Antigen LB39AA,Antigen SK29AA,MART1,MLANA,Melan A,MelanA,MelanA Protein)
- Atorvastatin Calcium(CI981,Lipitor)
- Tazarotene(AGN190168,Avage,Tazorac)
- Valproic Acid(Depakene,Stavzor,Valproate)
- 2,6 Diaminopurine(DAP)
- Peginterferon Alfa2a(Pegasys)
- ABT510
- Ixabepilone(mRNA2416)
- Tanespimycin(KOS953)
- PPI2458
- Ispinesib(CK0238273,SB715992)
- Pertuzumab(2C4, Omnitarg, Perjeta, RO4368451, rhuMAb2C4)
- OSI7904L(GS7904L)
- Dasatinib(BMS354825, Sprycel)
- Lorvotuzumab Mertansine(BB10901,IMGN901,huN901DM1)
- Anakinra(Kinaret,rIL1ra,rIL1RN)
- Axitinib(AG013736,Inlyta)
- Recombinant Human Endostatin(Endostar,rhEndostatin)
- Cinacalcet Hydrochloride(Mimpara,Sensipar)
- Motesanib Diphosphate(AMG 706)
- Cediranib Maleate(AZD2171,AZD2171 Maleate,Recentin)
- Bardoxolone(CDDO,RTA401)
- Degarelix(FE200486,Firmagon)
- Everolimus(Afinitor,Certican,RAD001,Votubia,Zortress)
- Retaspimycin Hydrochloride(IPI50)
- Lestaurtinib(CEP701,KT5555,SPM924)
- Tandutinib(CT53518,MLN518)
- Lucatumumab(CHIR1212,HCD122)
- Paclitaxel Loaded Polymeric Micelle(Cynviloq TM,Genexol PM,IG001)
- TPI287
- Volociximab(M200)
- Lenvatinib Mesylate(E7080,Lenvima)
- Belinostat(Beleodaq,PXD101)
- gp100:209217 210M Peptide Vaccine
- Ridaforolimus(AP23573,Deforolimus,MK 8669)
- Tremelimumab(ticilimumab, CP675, CP-675206)
- Bavituximab(Tarvacin)
- XL820
- BI2536
- Neratinib(HKI272,PB272,Nerlynx)
- MK0752
- Palbociclib(Ibrance,PD0332991)
- AG024322
- Atiprimod(SK&F106615,azaspirane)
- Figitumumab(CP751871)
- GMK562 Cell Vaccine
- Carfilzomib(Kyprolis,PR171)
- Brivanib Alaninate(BMS582664)
- Dacomitinib(PF00299804, PF0029980403, Vizimpro)
- Veliparib(ABT888)

- Pelareorep(POBB0209,Reolysin,Reovirus Serotype 3,Wildtype Reovirus)
- Pazopanib Hydrochloride(GW786034B,Votrient)
- Bosutinib(SKI606, Bosulif)
- Sagopilone(DE03757,Epothilone ZK219477,SHY03757A,ZKEPO,ZKEpothilone)
- 1018ISS(CPG1018)
- Oportuzumab Monatox(Proxinium,VB4845,Vicinium)
- BRaf VEGFR2 Inhibitor RAF265(CHIR265,RAF265)
- Talimogene Laherparepvec(Imlygic,OncoVEX GMCSF,TVEC)
- Tozasertib Lactate(L001281814,MK0457,VX680)
- Recombinant Thyrotropin Alfa(TSHalpha,Thyrogen)
- Mapatumumab(HGSETR1,HGSETR1,TRM 1 mAb)
- Bendamustine Hydrochloride(Bendeka,Cytostasan Hydrochloride,Levact,Ribomustin,SyB L0501,Treanda)
- Chloroquine
- Pegvisomant(Somavert,Trovert)
- Dacetuzumab(SGN40,huS2C6)
- Dovitinib Lactate(CHIR258,TKI258)
- Tivantinib(ARQ197)
- Plinabulin(NPI2358)
- Catumaxomab(Removab)
- Urelumab(BMS663513)
- Tesevatinib(EXEL7647,KD019,KD019,XL6 47)
- Barasertib(AZD1152,AZD2811)
- Mogamulizumab(KW-0761, KM8761,Poteligeo)
- CDK Inhibitor SNS032(BMS387032)
- Vintafolide(EC145)
- Iniparib(BSI201,SAR240550)
- Blinatumomab(Blincyto,MEDI538,MT103)
- CRLX101(IT101)
- Tertomotide(GV1001,PrimoVax)
- PM00104(Zalypsis)
- Obatoclax Mesylate(GX15070MS)
- Seliciclib(CYC202,R-roscovitine)
- Pegdinetanib(Angiocept,BMS844203,CT3 22)
- Tasisulam(LY573636)
- Fresolimumab(GC1008)
- Asparaginase Erwinia chrysanthemi(Crisantaspasum,Cristantas pase,Erwinase,Erwinaze)
- ZYC300
- Marizomib,(ML858,NPI0052,Salinospora mide A)
- Crenolanib(CP868596)
- Efatutazone Dihydrochloride(CS7017,Inolitazone Dihydrochloride)
- E2F1 Pathway Activator ARQ171
- Vemurafenib(PLX4032,RG7204,RO51854 26,Zelboraf)
- Tagraxofusp-erzs(SL-401, Tagraxofusp, Elzonris, DT388IL3 fusion protein)
- Navitoclax(A855071.0,ABT263)
- Auranofin(Ridaura)
- Gimatecan(LBQ707, ST1481)
- Oblimersen Sodium(Augmerosen,G3139,Genasense)
- Lapatinib Ditosylate(Tykerb)
- Cabazitaxel(Jevtana, RPR116258A,Taxoid XRP6258)
- Brentuximab Vedotin(ADC SGN35, Adcetris, SGN35)
- Panobinostat(Faridak,Farydak,LBH589)
- Ofatumumab(GSK1841157, Arzerra)
- AZD7762
- Elotuzumab(BMS901608,Empliciti,HuLuc 63,PDL063)
- Rilotumumab(AMG102)
- Liposome encapsulated Daunorubicin Cytarabine(CPX351,Liposomal AraC Daunorubicin,Vyxeos)
- Nivolumab(BMS936558,MDX1106,NIVO, ONO4538,Opdivo)
- Moxetumomab Pasudotox(HA22, CAT8015, GCR8015, Lumoxiti)
- Abiraterone Acetate(CB7630,Yonsa,Zytiga)
- Abexinostat(CRA024781,PCI24781)

- Aldoxorubicin(DOXOEMCH,DoxorubicinEMCH,INNO206)
- Cobimetinib(Cotellic,GDC0973,XL518)
- Quizartinib(AC010220,AC220)
- Rindopepimut(CDX110,PF04948568)
- PRLX93936
- Obinutuzumab(Afutuzumab,GA101,R7159,RO5072759, Gazyva)
- Ramucirumab(Cyramza,IMC1121B,LY3009806)
- OSI930
- GCS100
- Pidilizumab(CT011,MDV9300)
- Conatumumab(AMG 655,XG1048)
- Anti PSCA Monoclonal Antibody AGS1C4D4(MK4721)
- AVE9633
- Ganitumab(AMG479)
- Inotuzumab Ozogamicin(CMC544,WAY207294,Besponsa)
- Pioglitazone
- Voxtalisib(SAR245409,XL765)
- Pilaralisib(SAR245408,XL147)
- Alisertib(MLN8237)
- Olaparib(AZD2281,KU0059436,Lynparza)
- Apatorsen(ISIS306053,OGX427)
- GS9219
- Enzalutamide(ASP9785, MDV3100,Xtandi)
- Linifanib(ABT869)
- hTERT I540 R572Y D988Y Multipeptide Vaccine
- Pomalidomide(Actimid,CC4047,Imnovid,Pomalyst)
- Bazedoxifene(TSE424,WAY140424)
- AntiCD19 Monoclonal Antibody MDX1342
- Daratumumab(Darzalex,HuMaxCD38,JNJ54767414)
- Apatinib(YN968D1)
- Zoptarelin Doxorubicin(AEZS108,AN152,ZEN008)
- Vismodegib(Erivedge,GDC0449)
- Crizotinib(PF2341066,Xalkori)
- BGT226
- Lanreotide Acetate(Somatuline Depot)
- Corticorelin Acetate(Xerecept,hCRF)
- Tasidotin(ILX651)
- Histrelin Acetate(Supprelin, Vantas)
- Ganetespib(STA9090)
- Glesatinib(MG90265, MG90265X)
- Tucatinib(ARRY380, Irbinitinib, ONT380)
- Pevonedistat(MLN4924)
- Trametinib(GSK1120212, JTP74057,Mekinist)
- ENMD 2076
- Ponatinib Hydrochloride(Iclusig)
- Calaspargase Pegol(EZN2285,Oncaspar IV, SCPEG E Coli L Asparaginase)
- Regorafenib(BAY734506,Stivarga)
- PGG BetaGlucan(Imprime PGG)
- IPH2101
- PF04217903
- Idelalisib(Zydelig,CAL101, GS1101)
- Pacritinib(SB1518)
- BT062
- Patritumab(AMG888, U31287)
- Dinaciclib(MK7965,SCH727965)
- Ammonia N13
- Olaratumab(IMC3G3,Lartruvo)
- Cixutumumab(AntiIGF1R Recombinant MOAB IMCA12)
- Elesclomol Sodium(STA4783)
- Talmapimod(SCIO469)
- BHQ880
- Codrituzumab(GC33)
- Foretinib(GSK1363089,XL880)
- Niraparib(Zejula, MK4827)
- Patidegib(FIN5,IP9 Free Base,IPI926,IPI926 Free Base,Saridegib)
- Talotrexin Ammonium(PT523)
- Dacinostat(NVPLAQ824)
- Ibrutinib(CRA032765,Imbruvica,PCI32765)
- Alpha1 Proteinase Inhibitor Human(A1AT,A1PI,AAT,Alpha 1 Antitrypsin, Aralast, ProlastinC)
- Cold Contaminant-free Iobenguane I 131(Azedra, Ultratrace MIBG)
- GDC0941 Bismesylate
- Lurbinectedin(PM01183)
- PF03084014
- Sonidegib(Erismodegib,LDE225,Odomzo)

- Dabrafenib(Tafinlar, GSK2118436A, GSK2118436)
- AntiKSP AntiVEGF siRNAs ALNVSP02
- BMS863233
- Trastuzumab Emtansine(ADO TRASTUZUMAB EMTANSINE, Kadcyla, PRO132365, RO5304020, TDM1 ,Trastuzumab DM1,Trastuzumab MCCDM1)
- Ixazomib Citrate(MLN9708,Ninlaro)
- TAK901
- MEDI573
- Tovetumab(MEDI575)
- BCG Solution(Bacillus Calmette Guerin Solution,TICE BCG Solution)
- Tocilizumab(Actemra,MRA,R1569)
- Imetelstat Sodium
- Glasdegib(PF04449913, Daurismo)
- Binimetinib(ARRY162,ARRY438162,MEK162,Mektovi)
- Uprosertib(GSK2141795)
- Tivozanib(AV951)
- Onalespib(AT13387)
- Smac Mimetic GDC0152
- Necitumumab(IMC11F8,Portrazza)
- Infigratinib(BGJ398)
- Pexidartinib(PLX3397)
- CXCR4 Antagonist BL8040(BKT140)
- Momelotinib(CYT387,GS0387)
- Vistusertib(AZD2014)
- Sapanisertib(INK128,MLN0128,TAK228)
- Capmatinib(INC280,INCB28060,INCB028060)
- Buparlisib(BKM120)
- Autologous Melanoma Lysate Pulsed Dendritic Cell Vaccine
- XL019
- MK2206
- Smac Mimetic LCL161
- Mitoguazone Dihydrochloride(MGBG 2HCl)
- Urokinase Derived Peptide A6
- Trebananib(AMG386)
- AS1411
- MDM2 Antagonist RO5045337(R7112)
- Adavosertib(AZD1775,MK1775)
- Tabalumab(LY2127399)
- Gedatolisib(PKI587, PF05212384)
- Apalutamide(ARN509,Erleada,JNJ56021927)
- Epacadostat(INCDB024360)
- CP547632
- Alpelisib(BYL719, Piqray)
- AntiCD30 CD16A Monoclonal Antibody AFM13
- Lutetium Lu 177 Dotatate(Lutathera)
- Vosaroxin(AG7352,SPC595)
- AMG337
- Nilotinib Hydrochloride Monohydrate(AMN107, Tasigna)
- Ribociclib(LEE011, Kisqali)
- MEDI3617
- Anti PSMA Monoclonal Antibody MMAE Conjugate
- PI3K Inhibitor ZSTK474
- Merestinib(LY2801653)
- LY2875358
- Talazoparib(BMN673, Talzenna)
- Ulocuplumab(BMS936564,MDX1338)
- Teprotumumab(R1507,RO4858696,RV001)
- Vedolizumab(Entyvio,LDP02,MLN0002,MLN02)
- ARRY382
- Ricolinostat(ACY1215)
- LFA102
- Afatinib Dimaleate(BIBW 2992MA2,Gilotrif)
- PF04136309
- Abemaciclib(LY2835219,Verzenio)
- PWT33597 Mesylate
- AGS22M6E
- Ruxolitinib Phosphate(INCB18424 Phosphate,Jakafi)
- Cabozantinib Smalate(Cabometyx,Cometriq,XL184)
- NS018
- Encorafenib(Braftovi,LGX818)
- Brigatinib(AP26113,Alunbrig)
- AS703988 MSC2015103B
- Varlilumab(CDX1127)
- Duvelisib(Copiktra,INK1197,IPI145)
- Paclitaxel Trevatide(ANG1005,GRN1005)
- Rociletinib(CO1686)

- IMGN529
- Pegylated Recom Lasparaginase Erwinia chrysanthemi(Asparec)
- MLN0264
- Molibresib(GSK525762)
- Alectinib(AF802,Alecensa,CH5424802,RG 7853,RO5424802)
- Olmutinib(BI1482694,HM61713)
- Selinexor(KPT-330,Xpovio)
- Trifluridine and Tipiracil Hydrochloride(Lonsurf,TAS102)
- Ensartinib(X396)
- Tisagenlecleucel(CART-19,CTL019,Kymriah)
- Thioureidobutyronitrile(Kevetrin)
- Venetoclax(ABT0199,ABT199,GDC0199,R G7601,Venclexta)
- Durvalumab(MEDI4736, Imfinzi)
- Erdafitinib(JNJ-42756493, Balversa)
- Lirilumab(BMS986015,IPH2102)
- PVX410
- Polatuzumab Vedotin(RG7596,DCDS4501A,FCU 2711,Polivy)
- Umbralisib(RP5264,TGR12020)
- Ulixertinib(BVD523,VRT752271)
- Amcasertib(BBI503)
- Darolutamide(BAY1841788,ODM201,Nu beqa)
- Brilanestrant(ARN810,GDC0810,RO7056 118)
- Atezolizumab(MPDL3280A,RG7446,RO55 41267,Tecentriq)
- Pembrolizumab(Keytruda,Lambrolizuma b,MK3475,SCH900475)
- GDC0994
- Tazemetostat(Tazverik,E7438,EPZ6438)
- Vadastuximab Talirine(SGNCD33A)
- Enasidenib(AG221,CC90007,Idhifa)
- AKT 1 2 Inhibitor BAY1125976(AY1125976)
- CPI0610
- Relatlimab(BMS986016)
- BET Inhibitor RO6870810(RG6146,TEN010)
- TAK659
- Acalabrutinib(ACP196,Calquence)
- Lorlatinib(PF06463922, Lorbrena)
- UAE Inhibitor TAK243(AOB87172,MLN7243)
- Akt ERK Inhibitor ONC201(TIC10)
- AntiBCMA Conjugate GSK2857916(J6M0mcMMAF)
- Ivosidenib(AG120, Tibsovo)
- Asciminib(ABL001)
- Enfortumab Vedotin(Padcev, ASG22CE)
- TLR7 8 9 Antagonist IMO8400
- Entrectinib(RXDX101,Rozlytrek)
- Nazartinib(EGF816)
- Naquotinib(ASP8273)
- Ceritinib(LDK378, Zykadia)
- Larotrectinib(LOXO 101, ARRY470)
- Osimertinib(AZD9291, Mereletinib, Tagrisso)
- Utomilumab(PF5082566, PF2566)
- Gilteritinib(ASP2215, Xospata)
- DLYE5953A
- Pegylated Liposomal Nanoparticle Docetaxel Prodrug MNK010(MP35491)
- Avelumab(MSB0010718C, Bavencio)
- Taselisib(GDC0032,RO5537381)
- Oral Azacitidine(CC486)
- Altiratinib(DCC2701)
- MGD007
- STM434
- REGN1979
- RO6958688
- Mavelertinib(PF06747775)
- Axicabtagene Ciloleucel(KTEC19, Yescarta)
- CC90002
- VLX1570
- Cemiplimab(REGN2810)
- Spartalizumab(PDR001)
- Citarinostat(ACY241,CC96241,HDACIN2)
- PLX9486
- AntiFGFR3 Monoclonal Antibody B701
- Avapritinib(Ayvakit,BLU285)
- ASTX660
- SC003
- ABBV085
- Bintrafusp Alfa(MSB0011359C,M7824)
- Pegzilarginase(AEB1102,CoArgIPEG)

- Topical Fluorouracil(ActinoHermal, Arumel, Carac, Cytosafe, Efudex, Efurix, Fiverocil, Fluoroplex, Flurox, Timazin, Tolak)
- KTN0158
- IMGN779
- MIW815(ADUS100)
- GS5829
- INCAGN01876
- Apilimod Dimesylate Capsule(LAM 002A)
- AP32788
- TSR042
- AntiGITR Monoclonal Antibody GWN323
- Defactinib Hydrochloride(PF04554878, VS6063)
- LY3300054
- Trastuzumab Deruxtecan(DS8201a,Enhertu,WHO10516)
- H3B8800
- ATR Kinase Inhibitor VX803
- Aldesleukin Prodrug NKTR214
- Cereblon Modulator CC90009
- Rucaparib Camsylate(C0338, Rubraca, Rucaparib Phosphate)
- LY3214996
- PEN221
- Albumin binding Cisplatin Prodrug BTP114
- XMT1522
- BMS986179
- Vecabrutinib(BIIB062,BSK4841,FP182,SNS062)
- LOXO195
- Rituximab and Hyaluronidase Human(Rituxan Hycela)
- Copanlisib Hydrochloride(BAY806946,Aliqopa)
- Icotinib(BPI2009)
- Zanubrutinib(Brukinsa,BGD3111)
- INCB001158(CB1158)
- PF06863135
- SAR439459
- AZD1390
- Anti-TROP2/DXd Antibody-drug Conjugate DS-1062a
- Allogeneic GMCSF secreting Lethally Irradiated Pancreatic Tumor Cell Vaccine
- ASTX029
- DHES0815A
- Other NOS
- Other antineoplastic
- Other hormone
- Clinical Trial Drug not specified

Appendix 3. Curation Instructions for Medical Oncologist Assessments

- Find the date of diagnosis of the cancer of interest.
- Some patients have more than one cancer diagnosis. Do review and curate only notes for the cancer of interest.
- Do review and curate one clinical assessment per month, beginning at time of diagnosis (+/- 30 days).
  - Do choose first Medical Oncology note of the month that is authored by an MD actively following the patient for the cancer of interest.
  - If there is no note by a medical oncologist (MD) that month. Do use the first note by a nurse practitioner or physician assistant (NP/PA) from a medical oncology practice.
    - Rarely look at more than one note per month. If the patient has imaging scans after the first note in the month, look on or up to 7 days following the imaging scans.
  - Do give priority to internal visit notes. Only curate notes from an outside institution if there is no DFCI/Partners oncology note in that month. Prioritize as follows:
    1. Internal MD
    2. Internal NP/PA
    3. External MD
    4. External NP/PA
  - If a fellow has a note and an attending physician adds an addendum, review any information in the Summary/impression/Evaluation/Plan, including the addendum.
  - Do not use notes from Radiation Oncology. Surgery/Surgical Oncology, inpatient care, primary care, or other specialists not related to cancer (e.g. dermatology, cardiology).
    - Patients with early stage cancers are sometimes only followed by a surgical oncologist. This may mean that they are many months without notes that qualify for curation. That is ok!
  - Rarely curate notes by an oncologist from a different specialty than the cancer of interest. Occasionally a patient will transfer care to a different type of oncologist due to the particulars of their disease. For example. a patient with brain metastases from their primary cancer may be followed by Neuro Oncology. Do curate these notes if notes from the primary oncologist are not available. Do not curate notes from a different oncology specialist that are pertaining to a different cancer diagnosis.
- Do review only the Impression/Plan section at the bottom of the note as well as the reason for visit.
  - Rarely a medical oncology provider will summarize the cancer status directly above the Impression/Plan section, which can be reviewed for curation.
- Do not review any of the other sections. including the physical exam, interval history, lab results. etc.
  - The Impression/Plan section may have a different name in medical oncologist progress notes; other section headers could include Assessment, Summary, Conclusion, Problem List Items Addressed this Visit.
  - When there are no section headings in a provider's note. Do review everything located beneath the physical exam.

Appendix 4. NSCLC Distant Metastasis Classification

ICD-O-3 Topography Code	Metastasis Classification
C00.0 External upper lip	Distant (Head and neck)
C00.1 External lower lip	Distant (Head and neck)
C00.2 External lip NOS	Distant (Head and neck)
C00.3 Mucosa of upper lip	Distant (Head and neck)
C00.4 Mucosa of lower lip	Distant (Head and neck)
C00.5 Mucosa of lip NOS	Distant (Head and neck)
C00.6 Commissure of lip	Distant (Head and neck)
C00.8 Overlapping lesion of lip	Distant (Head and neck)
C00.9 Lip NOS	Distant (Head and neck)
C01.9 Base of tongue NOS	Distant (Head and neck)
C02.0 Dorsal surface of tongue NOS	Distant (Head and neck)
C02.1 Border of tongue	Distant (Head and neck)
C02.2 Ventral surface of tongue NOS	Distant (Head and neck)
C02.3 Anterior 2/3 of tongue NOS	Distant (Head and neck)
C02.4 Lingual tonsil	Distant (Head and neck)
C02.8 Overlapping lesion of tongue	Distant (Head and neck)
C02.9 Tongue NOS	Distant (Head and neck)
C03.0 Upper Gum	Distant (Head and neck)
C03.1 Lower gum	Distant (Head and neck)
C03.9 Gum NOS	Distant (Head and neck)
C04.0 Anterior floor of mouth	Distant (Head and neck)
C04.1 Lateral floor of mouth	Distant (Head and neck)
C04.8 Overlapping lesion of floor of mouth	Distant (Head and neck)
C04.9 Floor of mouth NOS	Distant (Head and neck)
C05.0 Hard palate	Distant (Head and neck)
C05.1 Soft palate NOS	Distant (Head and neck)
C05.2 Uvula	Distant (Head and neck)
C05.8 Overlapping lesion of palate	Distant (Head and neck)
C05.9 Palate NOS	Distant (Head and neck)
C06.0 Cheek mucosa	Distant (Head and neck)
C06.1 Vestibule of mouth	Distant (Head and neck)
C06.2 Retromolar area	Distant (Head and neck)
C06.8 Overlapping lesion of other and unspecified parts of mouth	Distant (Head and neck)
C06.9 Mouth NOS	Distant (Head and neck)
C07.9 Parotid gland	Distant (Head and neck)
C08.0 Submandibular gland	Distant (Head and neck)
C08.1 Sublingual gland	Distant (Head and neck)
C08.8 Overlapping lesion of major salivary glands	Distant (Head and neck)
C08.9 Major salivary gland NOS	Distant (Head and neck)
C09.0 Tonsillar fossa	Distant (Head and neck)
C09.1 Tonsillar pillar	Distant (Head and neck)
C09.8 Overlapping lesion of tonsil	Distant (Head and neck)
C09.9 Tonsil NOS	Distant (Head and neck)
C10.0 Vallecula	Distant (Head and neck)
C10.1 Anterior surface of epiglottis	Distant (Head and neck)
C10.2 Lateral wall of oropharynx	Distant (Head and neck)
C10.3 Posterior wall of oropharynx	Distant (Head and neck)
C10.4 Branchial cleft	Distant (Head and neck)
C10.8 Overlapping lesions of oropharynx	Distant (Head and neck)
C10.9 Oropharynx NOS	Distant (Head and neck)
C11.0 Superior wall of nasopharynx	Distant (Head and neck)
C11.1 Posterior wall of nasopharynx	Distant (Head and neck)

ICD-O-3 Topography Code	Metastasis Classification
C11.2 Lateral wall of nasopharynx	Distant (Head and neck)
C11.3 Anterior wall of nasopharynx	Distant (Head and neck)
C11.8 Overlapping lesion of nasopharynx	Distant (Head and neck)
C11.9 Nasopharynx NOS	Distant (Head and neck)
C12.9 Pyriform sinus	Distant (Head and neck)
C13.0 Postcricoid region	Distant (Head and neck)
C13.1 Hypopharyngeal aspect of aryepiglottic fold	Distant (Head and neck)
C13.2 Posterior wall of hypopharynx	Distant (Head and neck)
C13.8 Overlapping lesion of hypopharynx	Distant (Head and neck)
C13.9 Hypopharynx NOS	Distant (Head and neck)
C14.0 Pharynx NOS	Distant (Head and neck)
C14.2 Waldeyer ring	Distant (Head and neck)
C14.8 Overlapping lesion of lip oral cavity and pharynx	Distant (Head and neck)
C15.0 Cervical esophagus	Distant (Head and neck)
C15.1 Thoracic esophagus	Distant (Thorax)
C15.2 Abdominal esophagus	Distant (Abdomen)
C15.3 Upper third of esophagus	Distant (Head and neck)
C15.4 Middle third of esophagus	Distant (Thorax)
C15.5 Lower third of esophagus	Distant (Abdomen)
C15.8 Overlapping lesion of esophagus	Distant (Rare and nos)
C15.9 Esophagus NOS	Distant (Rare and nos)
C16.0 Cardia NOS	Distant (Abdomen)
C16.1 Fundus of stomach	Distant (Abdomen)
C16.2 Body of stomach	Distant (Abdomen)
C16.3 Gastric antrum	Distant (Abdomen)
C16.4 Pylorus	Distant (Abdomen)
C16.5 Lesser curvature of stomach NOS	Distant (Abdomen)
C16.6 Greater curvature of stomach NOS	Distant (Abdomen)
C16.8 Overlapping lesion of stomach	Distant (Abdomen)
C16.9 Stomach NOS	Distant (Abdomen)
C17.0 Duodenum	Distant (Abdomen)
C17.1 Jejunum	Distant (Abdomen)
C17.2 Ileum	Distant (Abdomen)
C17.3 Meckel diverticulum	Distant (Abdomen)
C17.8 Overlapping lesion of small intestine	Distant (Abdomen)
C17.9 Small intestine NOS	Distant (Abdomen)
C18.0 Cecum	Distant (Abdomen)
C18.1 Appendix	Distant (Abdomen)
C18.2 Ascending colon	Distant (Abdomen)
C18.3 Hepatic flexure of colon	Distant (Abdomen)
C18.4 Transverse colon	Distant (Abdomen)
C18.5 Splenic flexure of colon	Distant (Abdomen)
C18.6 Descending colon	Distant (Abdomen)
C18.7 Sigmoid colon	Distant (Pelvis)
C18.8 Overlapping lesion of colon	Distant (Abdomen)
C18.9 Colon NOS	Distant (Abdomen)
C19.9 Rectosigmoid junction	Distant (Pelvis)
C20.9 Rectum NOS	Distant (Pelvis)
C21.0 Anus NOS	Distant (Pelvis)
C21.1 Anal canal	Distant (Pelvis)
C21.2 Cloacogenic zone	Distant (Pelvis)
C21.8 Overlapping lesion of rectum anus and anal canal	Distant (Pelvis)
C22.0 Liver	Distant (Liver)
C22.1 intrahepatic bile duct	Distant (Abdomen)

ICD-O-3 Topography Code	Metastasis Classification
C23.9 Gallbladder	Distant (Abdomen)
C24.0 Extrahepatic bile duct	Distant (Abdomen)
C24.1 Ampulla of Vater	Distant (Abdomen)
C24.8 Overlapping lesion of biliary tract	Distant (Abdomen)
C24.9 Biliary tract NOS	Distant (Abdomen)
C25.0 Head of pancreas	Distant (Abdomen)
C25.1 Body of pancreas	Distant (Abdomen)
C25.2 Tail of pancreas	Distant (Abdomen)
C25.3 Pancreatic duct	Distant (Abdomen)
C25.4 Islets of Langerhans	Distant (Abdomen)
C25.7 Other specified parts of pancreas	Distant (Abdomen)
C25.8 Overlapping lesion of pancreas	Distant (Abdomen)
C25.9 Pancreas NOS	Distant (Abdomen)
C26.0 Intestinal tract NOS	Distant (Abdomen)
C26.8 Overlapping lesion of digestive system	Distant (Abdomen)
C26.9 Gastrointestinal tract NOS	Distant (Abdomen)
C30.0 Nasal cavity	Distant (Head and neck)
C30.1 Middle ear	Distant (Head and neck)
C31.0 Maxillary sinus	Distant (Head and neck)
C31.1 Ethmoid sinus	Distant (Head and neck)
C31.2 Frontal sinus	Distant (Head and neck)
C31.3 Sphenoid sinus	Distant (Head and neck)
C31.8 Overlapping lesion of accessory sinuses	Distant (Head and neck)
C31.9 Accessory sinus NOS	Distant (Head and neck)
C32.0 Glottis	Distant (Head and neck)
C32.1 Supraglottis	Distant (Head and neck)
C32.2 Subglottis	Distant (Head and neck)
C32.3 Laryngeal cartilage	Distant (Head and neck)
C32.8 Overlapping lesion of larynx	Distant (Head and neck)
C32.9 Larynx NOS	Distant (Head and neck)
C33.9 Trachea	Distant (Thorax)
C34.0 Main bronchus	Local/Regional (Thorax)
C34.1 Upper lobe lung	Local/Regional (Thorax)
C34.2 Middle lobe lung	Local/Regional (Thorax)
C34.3 Lower lobe lung	Local/Regional (Thorax)
C34.8 Overlapping lesion of lung	Local/Regional (Thorax)
C34.9 Lung NOS	Local/Regional (Thorax)
C37.9 Thymus	Distant (Thorax)
C38.0 Heart	Distant (Thorax)
C38.1 Anterior mediastinum	Distant (Thorax)
C38.2 Posterior mediastinum	Distant (Thorax)
C38.3 Mediastinum NOS	Distant (Thorax)
C38.4 Pleura NOS	Distant (Thorax)
C38.8 Overlapping lesion of heart mediastinum and pleura	Distant (Thorax)
C39.0 Upper respiratory tract NOS	Distant (Thorax)
C39.8 Overlapping lesion of respiratory system and intrathoracic organs	Distant (Thorax)
C39.9 Ill-defined sites within respiratory system	Distant (Thorax)
C40.0 Long bones of upper limb scapula and associated joints	Distant (Bone)
C40.1 Short bones of upper limb and associated joints	Distant (Bone)
C40.2 Long bones of lower limb and associated joints	Distant (Bone)
C40.3 Short bones of lower limb and associated joints	Distant (Bone)
C40.8 Overlapping lesion of bones joints and articular cartilage of limbs	Distant (Bone)
C40.9 Bone of limb NOS	Distant (Bone)

ICD-O-3 Topography Code	Metastasis Classification
C41.0 Bones of skull and face and associated joints	Distant (Bone)
C41.1 Mandible	Distant (Head and neck)
C41.2 Vertebral column	Distant (Bone)
C41.3 Rib sternum clavicle and associated joints	Distant (Thorax)
C41.4 Pelvic bones sacrum coccyx and associated joints	Distant (Bone)
C41.8 Overlapping lesion of bones joints and articular cartilage	Distant (Bone)
C41.9 Bone NOS	Distant (Bone)
C42.2 Spleen	Distant (Abdomen)
C44.0 Skin of lip NOS	Distant (Head and neck)
C44.1 Eyelid	Distant (Head and neck)
C44.2 External ear	Distant (Head and neck)
C44.3 Skin of other and unspecified parts of face	Distant (Head and neck)
C44.4 Skin of scalp and neck	Distant (Head and neck)
C44.6 Skin of upper limb and shoulder	Distant (Extremity)
C44.7 Skin of lower limb and hip	Distant (Extremity)
C44.8 Overlapping lesion of skin	Distant (Rare and nos)
C44.9 Skin NOS	Distant (Rare and nos)
C47.0 Peripheral nerves and autonomic nervous system of head face and neck	Distant (Head and neck)
C47.1 Peripheral nerves and autonomic nervous system of upper limb and shoulder	Distant (Extremity)
C47.2 Peripheral nerves and autonomic nervous system of lower limb and hip	Distant (Extremity)
C47.3 Peripheral nerves and autonomic nervous system of thorax	Local/Regional (Thorax)
C47.4 Peripheral nerves and autonomic nervous system of abdomen	Distant (Abdomen)
C47.5 Peripheral nerves and autonomic nervous system of pelvis	Distant (Pelvis)
C47.6 Peripheral nerves and autonomic nervous system of trunk NOS	Distant (Abdomen)
C47.8 Overlapping lesion of peripheral nerves and autonomic nervous system	Distant (Rare and nos)
C47.9 Autonomic nervous system NOS	Distant (Rare and nos)
C48.0 Retroperitoneum	Distant (Abdomen)
C48.1 Specified parts of peritoneum	Distant (Abdomen)
C48.2 Peritoneum NOS	Distant (Abdomen)
C48.8 Overlapping lesion of retroperitoneum and peritoneum	Distant (Abdomen)
C49.0 Connective Subcutaneous and other soft tissues of head face and neck	Distant (Head and neck)
C49.1 Connective Subcutaneous and other soft tissues of upper limb and shoulder	Distant (Extremity)
C49.2 Connective Subcutaneous and other soft tissues of lower limb and hip	Distant (Extremity)
C49.3 Connective Subcutaneous and other soft tissues of thorax	Local/Regional (Thorax)
C49.4 Connective Subcutaneous and other soft tissues of abdomen	Distant (Abdomen)
C49.5 Connective Subcutaneous and other soft tissues of pelvis	Distant (Pelvis)
C49.6 Connective Subcutaneous and other soft tissues of trunk NOS	Distant (Abdomen)
C49.8 Overlapping lesion of connective subcutaneous and other soft tissues	Distant (Rare and nos)
C49.9 Connective Subcutaneous and other soft tissues NOS	Distant (Rare and nos)

ICD-O-3 Topography Code	Metastasis Classification
C50.0 Nipple	Distant (Breast)
C50.1 Central portion of breast	Distant (Breast)
C50.2 Upper-inner quadrant of breast	Distant (Breast)
C50.3 Lower-inner quadrant of breast	Distant (Breast)
C50.4 Upper-outer quadrant of breast	Distant (Breast)
C50.5 Lower-outer quadrant of breast	Distant (Breast)
C50.6 Axillary tail of breast	Distant (Breast)
C50.8 Overlapping lesion of breast	Distant (Breast)
C50.9 Breast NOS	Distant (Breast)
C51.0 Labium majus	Distant (Pelvis)
C51.1 Labium minus	Distant (Pelvis)
C51.2 Clitoris	Distant (Pelvis)
C51.8 Overlapping lesion of vulva	Distant (Pelvis)
C51.9 Vulva NOS	Distant (Pelvis)
C52.9 Vagina NOS	Distant (Pelvis)
C53.0 Endocervix	Distant (Pelvis)
C53.1 Exocervix	Distant (Pelvis)
C53.8 Overlapping lesion of cervix uteri	Distant (Pelvis)
C53.9 Cervix uteri	Distant (Pelvis)
C54.0 Isthmus uteri	Distant (Pelvis)
C54.1 Endometrium	Distant (Pelvis)
C54.2 Myometrium	Distant (Pelvis)
C54.3 Fundus uteri	Distant (Pelvis)
C54.8 Overlapping lesion of corpus uteri	Distant (Pelvis)
C54.9 Corpus uteri	Distant (Pelvis)
C55.9 Uterus NOS	Distant (Pelvis)
C56.9 Ovary	Distant (Pelvis)
C57.0 Fallopian tube	Distant (Pelvis)
C57.1 Broad ligament	Distant (Pelvis)
C57.2 Round ligament	Distant (Pelvis)
C57.3 Parametrium	Distant (Pelvis)
C57.4 Uterine adnexa	Distant (Pelvis)
C57.7 Other specified parts of female genital organs	Distant (Pelvis)
C57.8 Overlapping lesion of female genital organs	Distant (Pelvis)
C57.9 Female genital tract NOS	Distant (Pelvis)
C58.9 Placenta	Distant (Pelvis)
C60.0 Prepuce	Distant (Pelvis)
C60.1 Glans penis	Distant (Pelvis)
C60.2 Body of penis	Distant (Pelvis)
C60.8 Overlapping lesion of penis	Distant (Pelvis)
C60.9 Penis NOS	Distant (Pelvis)
C61.9 Prostate gland	Distant (Pelvis)
C62.0 Undescended testis	Distant (Pelvis)
C62.1 Descended testis	Distant (Pelvis)
C62.9 Testis NOS	Distant (Pelvis)
C63.0 Epididymis	Distant (Pelvis)
C63.1 Spermatic cord	Distant (Pelvis)
C63.2 Scrotum NOS	Distant (Pelvis)
C63.7 Other specified parts of male genital organs	Distant (Pelvis)
C63.8 Overlapping lesion of male genital organs	Distant (Pelvis)
C63.9 Male genital organs NOS	Distant (Pelvis)
C64.9 Kidney NOS	Distant (Abdomen)
C65.9 Renal pelvis	Distant (Abdomen)
C66.9 Ureter	Distant (Pelvis)
C67.0 Trigone of bladder	Distant (Pelvis)
C67.1 Dome of bladder	Distant (Pelvis)

ICD-O-3 Topography Code	Metastasis Classification
C67.2 Lateral wall of bladder	Distant (Pelvis)
C67.3 Anterior wall of bladder	Distant (Pelvis)
C67.4 Posterior wall of bladder	Distant (Pelvis)
C67.5 Bladder neck	Distant (Pelvis)
C67.6 Ureteric orifice	Distant (Pelvis)
C67.7 Urachus	Distant (Pelvis)
C67.8 Overlapping lesion of bladder	Distant (Pelvis)
C67.9 Bladder NOS	Distant (Pelvis)
C68.0 Urethra	Distant (Pelvis)
C68.1 Paraurethral gland	Distant (Pelvis)
C68.8 Overlapping lesion of urinary organs	Distant (Pelvis)
C68.9 Urinary system NOS	Distant (Pelvis)
C69.0 Conjunctiva	Distant (Head and neck)
C69.1 Cornea NOS	Distant (Head and neck)
C69.2 Retina	Distant (Brain/cns)
C69.3 Choroid	Distant (Head and neck)
C69.4 Ciliary body	Distant (Head and neck)
C69.5 Lacrimal gland	Distant (Head and neck)
C69.6 Orbit NOS	Distant (Head and neck)
C69.8 Overlapping lesion of eye and adnexa	Distant (Head and neck)
C69.9 Eye NOS	Distant (Head and neck)
C70.0 Cerebral meninges	Distant (Brain/cns)
C70.1 Spinal meninges	Distant (Brain/cns)
C70.9 Meninges NOS	Distant (Brain/cns)
C71.0 Cerebrum	Distant (Brain/cns)
C71.1 Frontal lobe	Distant (Brain/cns)
C71.2 Temporal lobe	Distant (Brain/cns)
C71.3 Parietal lobe	Distant (Brain/cns)
C71.4 Occipital lobe	Distant (Brain/cns)
C71.5 Ventricle NOS	Distant (Brain/cns)
C71.6 Cerebellum NOS	Distant (Brain/cns)
C71.7 Brain stem	Distant (Brain/cns)
C71.8 Overlapping lesion of brain	Distant (Brain/cns)
C71.9 Brain NOS	Distant (Brain/cns)
C72.0 Spinal cord	Distant (Brain/cns)
C72.1 Cauda equina	Distant (Brain/cns)
C72.2 Olfactory nerve	Distant (Brain/cns)
C72.3 Optic nerve	Distant (Brain/cns)
C72.4 Acoustic nerve	Distant (Brain/cns)
C72.5 Cranial nerve NOS	Distant (Brain/cns)
C72.8 Overlapping lesion of brain and central nervous system	Distant (Brain/cns)
C72.9 Nervous system NOS	Distant (Rare and nos)
C73.9 Thyroid gland	Distant (Head and neck)
C74.0 Cortex of adrenal gland	Distant (Abdomen)
C74.1 Medulla of adrenal gland	Distant (Abdomen)
C74.9 Adrenal gland NOS	Distant (Abdomen)
C75.0 Parathyroid gland	Distant (Head and neck)
C75.1 Pituitary gland	Distant (Head and neck)
C75.2 Craniopharyngeal duct	Distant (Head and neck)
C75.3 Pineal gland	Distant (Head and neck)
C75.4 Carotid body	Distant (Head and neck)
C75.5 Aortic body and other paraganglia	Distant (Abdomen)
C75.8 Overlapping lesion of endocrine glands and related structures	Distant (Rare and nos)
C75.9 Endocrine gland NOS	Distant (Rare and nos)
C76.0 Head face or neck NOS	Distant (Head and neck)
C76.1 Thorax NOS	Local/Regional (Thorax)
C76.2 Abdomen NOS	Distant (Abdomen)
C76.3 Pelvis NOS	Distant (Pelvis)

ICD-O-3 Topography Code	Metastasis Classification
C76.4 Upper limb NOS	Distant (Extremity)
C76.5 Lower limb NOS	Distant (Extremity)
C76.7 Other ill-defined sites	Distant (Rare and nos)
C76.8 Overlapping lesion of ill-defined sites	Distant (Rare and nos)
C77.0 Lymph nodes of head face and neck	Distant (Head and neck)
C77.1 Intrathoracic lymph nodes	Local/Regional (Thorax)
C77.2 Intra-abdominal lymph nodes	Distant (Abdomen)
C77.3 Lymph nodes of axilla or arm	Distant (Thorax)
C77.4 Lymph nodes of inguinal region or leg	Distant (Extremity)
C77.5 Pelvic lymph nodes	Distant (Pelvis)
C77.8 Lymph nodes of multiple regions	Distant (Rare and nos)
C77.9 Lymph node NOS	
C80.9 Unknown primary site	
F10 Cerebral Spinal Fluid	Distant (Brain/cns)
F50 Pericardial Fluid	Distant (Thorax)
F20 Peritoneal Fluid/Ascites	Distant (Abdomen)
F30 Pleural Fluid	Distant (Thorax)
F40 Urine	Distant (Pelvis)