VENUSCANCER

Women's cancers: do variations in patterns of care explain the world-wide inequalities in survival and avoidable premature deaths?

Background, aim and specific objectives

Breast, ovarian and cervical cancers are a major public health problem world-wide. Every year, approximately 2.5 million women are diagnosed with one of these three cancers, and they account for over 900,000 deaths.[1] Many of these deaths are avoidable, even in low- and middle-income countries, where cancers in women represent a major economic burden, both to families that lose a mother and to the national economy.[2]

Reducing the number of cancer deaths in women will require improvements in prevention, but also more effective health systems, to improve the survival of women who do develop one of these cancers.[3] Yet access to safe surgery varies widely between the richest and poorest countries,[4] and in more than 30 of the poorest countries, radiotherapy services are not available at all.[5,6]

World-wide differences in survival from these cancers are striking.[7,8] Inequalities in survival also exist between and within high-income countries,[9] where the problem has been succinctly summarised: "political toleration of unfairness in access to affordable cancer treatment is unacceptable".[10]

The CONCORD programme established world-wide surveillance of trends in cancer survival in 2015 (CONCORD-2).[7] It documented for the first time the very wide global differences in survival trends for most major cancers. The third cycle of the CONCORD programme updated survival trends to 2014.[8] VENUSCANCER, embedded in the CONCORD programme, will examine in much greater depth why these enormous differences in five-year net survival still persist for women's cancers in 2010-2014 (breast cancer: from 66% in India to 91% in the United States; cervical cancer: 52% Ecuador *vs.* 77% Korea; ovarian cancer: 16% India *vs.* 57% Costa Rica). It will enable us to quantify the extent to which differences in stage at diagnosis or approaches to treatment can explain the differences in survival between high- and low-income countries.

The overall aim of this study is to provide actionable evidence for health policy that is designed to reduce the burden of women's cancers world-wide.

The two main objectives of VENUSCANCER are:

1) **to collect detailed data from the medical records of women diagnosed with breast, ovarian or cervical cancer in at least two countries per continent, in the most recent year during the period 2015-2018 for which data are available.**

This data call for VENUSCANCER relates to objective 1. Data will include the woman's age, residence, socio-economic status or educational level; the anatomic site, morphology, behaviour and grade of her tumour; the clinical investigations (stage at diagnosis and, where available, prognostic bio-markers); the type and date of treatment, and the woman's last known vital status (alive, dead, emigrated), with the corresponding date (see the following Data Specification)

The data will be analysed to examine whether inequalities in survival from women's cancers are attributable to differences in disease biology between populations, or to differences in patterns of care (e.g. stage at diagnosis, access to treatment), or to socio-economic status, or to broader aspects of society, such as national wealth, expenditure on health (% of GDP), or the status of universal health coverage (UHC).

2) **to estimate the number of avoidable premature deaths that are attributable to inequalities in five-year survival between and within countries.**

If your registry contributed data for CONCORD-3, we will include your data in the analyses related to objective 2, but please let us know if you would prefer not to be involved.

References

1. Ferlay J, Soerjomataram I, Dikshit R*, et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; **136**: E359-86. http://dx.doi.org/10.1002/ijc.29210
2. Ginsburg O, Bray F, Coleman MP*, et al.* The global burden of women's cancers: a grand challenge in global health [Lancet Series on Women's Cancers]. *Lancet* 2017; **389**: 847-60. http://dx.doi.org/10.1016/S0140-6736(16)31392-7
3. Coleman MP. Cancer survival: global surveillance will stimulate health policy and improve equity. *Lancet* 2014; **383**: 564-73. http://dx.doi.org/10.1016/S0140-6736(13)62225-4
4. Meara JG, Leather AJM, Hagander L*, et al.* Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; **386**: 569-624. http://dx.doi.org/10.1016/s0140-6736(15)60160-x
5. International Atomic Energy Agency. Inequity in cancer care: a global perspective. IAEA Human Health Reports No. 3. Vienna: IAEA, 2011. http://www-pub.iaea.org/MTCD/Publications/PDF/Pub1471_web.pdf (accessed 1 Dec 2019).
6. Atun R, Jaffray DA, Barton MB*, et al.* Expanding global access to radiotherapy. *Lancet Oncol* 2015; **16**: 1153-86. http://dx.doi.org/10.1016/S1470-2045(15)00222-3
7. Allemani C, Weir HK, Carreira H*, et al.* Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet* 2015; **385**: 977–1010. https://doi.org/10.1016/S0140-6736(14)62038-9
8. Allemani C, Matsuda T, Di Carlo V*, et al.* Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 2018; **391**: 1023-75. http://dx.doi.org/10.1016/S0140-6736(17)33326-3
9. Coleman MP, Forman D, Bryant H*, et al.* Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. *Lancet* 2011; **377**: 127-38.
10. Sullivan R, Peppercorn J, Sikora K*, et al.* Delivering affordable cancer care in high-income countries. *Lancet Oncol* 2011; **12**: 933-80.

1. INTRODUCTION

1.1 This document provides a detailed description of the data that each cancer registry is asked to provide for VENUSCANCER. It has been developed from the data specification used for the third cycle of the CONCORD programme (CONCORD-3).[1]

1.2 Data for VENUSCANCER will be sent to the Cancer Survival Group at the London School of Hygiene and Tropical Medicine (LSHTM). Quality control and analyses will be performed at LSHTM. We may receive more than 300 data files. **This annex sets out the data structure required for us to manage these files efficiently.**

1.3 We define three index sites – breast (women), cervix, ovary. **Please select data for these cancers with the *topography code*.** Although we are interested in women's cancers, it will be of interest to monitor the frequency of breast and peritoneal cancers in men, as well as the proportion of records with unknown sex for each of the three cancers.

1.4 For simplicity, we will use the word "cancer" to refer to all primary malignant neoplasms.

1.5 We define index cancers (see page 13) as those that:

- occur at an index site
- were diagnosed in women normally resident in the territory covered by the registry, and who are thus included in the cancer incidence rates routinely reported by the registry
- were diagnosed during the **most recent year of incidence** chosen by the registry

1.6 We will focus on patterns of care and on short-term or mid-term survival for women diagnosed with an index cancer in the **most recent year of incidence** for which the registry has collected (or has access to) data on stage, treatment and bio-markers. Where available, we will also examine data on diagnostic procedures, socio-economic status and lifestyle.

Data submission

1.7 You are invited to submit data on all three index cancers, but that is not a requirement. Some registries only register selected cancers (e.g., breast cancer).

1.8 Please contribute data for only ***one*** year of incidence**, preferably the most recent year during 2015-2018 for which you have complete data**. We may consider data for an earlier year (e.g. during 2012-2014), but please discuss this with us before submitting your data.

1.9 When planning your data submission, please ensure that (a) the incidence data are considered to be complete for the calendar year that you submit and (b) the follow-up of all patients for their vital status is also considered to be complete, at least up to 31 December of the year of incidence you submit, preferably 31 December of the most recent year available.

1.10 All analyses will be for adults (aged 15-99 years).

1.11 Analyses will include only ***invasive*** **primary malignancies (behaviour code 3)**. However, please submit **all tumour records for each index cancer**, including those that are benign (behaviour code 0), uncertain (1) or *in situ* (2). Data on tumour behaviour will enable comparison of the intensity of diagnostic activity between participating registries.

1.12 During the period 2015-2018, most registries will have used ICD-O-3[2] for coding tumour site, morphology and behaviour, in preference to ICD-9,[3] ICD-10[4] or ICD-O-2.[5] Most registries are now using the first revision of ICD-O-3, published in 2013.[6] Tumour site, topography and behaviour should be coded according to ICD-O-3. If your data are not coded to ICD-O-3, or its first revision, please discuss this with us before submitting your data.

1.13 **All cancer data files must have the same structure:**

- All data files will be checked for adherence to protocol, as the first step of quality control. Tables of protocol adherence will be sent to you shortly after the data submission. Data

files that do not meet the protocol cannot be used. If you are in doubt about how to construct your data files, please discuss this with us **before submitting your data.**

- **Every tumour record must have the same structure**. Please include all **core variables** and all **optional variables** in every record. Please do not leave any variables blank, or empty. For example, if you choose not to submit data for the optional variables on **socio-economic status** (variables 90-99) and **lifestyle** (variables 102-105), fill them with the appropriate "missing value" (see page 46), as follows:
  - *You do not collect a particular data item*. For example, you do not collect data on race/ethnicity (variable 7). Use the missing value for race/ethnicity (99) in **all tumour records.**
  - *No data are available for a given variable in a particular tumour record*. For example, a tumour record contains no data for behaviour (variable 21). Use the missing value for behaviour (9) in that tumour record.
  - *You decide not to supply a variable*. For example, you collect data on lifestyle, but you decide not to provide data on lifestyle. Use the missing value (see page 46) for all lifestyle variables (variables 102-105) in **all tumour records**.
- If it is routine practice in your registry to substitute an imputed value for a missing value (e.g. the month of the year), and some of the variables you submit contain imputed values, **please tell us**: we will ask you for a description of the imputation procedures.
- If tumour records in your database include a special code ("flag") to indicate when a missing value has been imputed, please discuss this with us **before submitting your data**. We will ask you for a description of how each flag has been generated.
- If you have modified a standard coding scheme (such as ICD-O-3) by adding special codes for local use in your registry, **please recode your data** to the standard form before submission. If you have any doubts about the appropriate conversion, please discuss this with us **before submitting your data**.
- Data files should be submitted using the **CONCORD File Transmission Utility (**CONCORD protocol **Annex 3)**. This is a safe, quick and convenient method that meets all data security requirements. It is free of charge for all participating cancer registries.

2. VARIABLE NAMES AND SHORT DESCRIPTIONS

2.1 The 59 **core variables** and the 52 **optional variables** in each record are shown below and on the next page. Details of the content and coding of each variable are given in Section 3: click on any variable name (**VARxx**) to jump to the relevant page.

Name	Short description
Demographics	
VAR1	Country
VAR2	Registry
VAR3	Person code
VAR4	Tumour code
VAR5	Sex
VAR6	Region *
VAR7	Race/ethnicity *
VAR8	Day of birth
VAR9	Month of birth
VAR10	Year of birth
VAR11	Day of diagnosis
VAR12	Month of diagnosis
VAR13	Year of diagnosis
Follow-up for vital status	
VAR14	Last known vital status
VAR15	Day of last known vital status
VAR16	Month of last known vital status
VAR17	Year of last known vital status
Tumour details	
VAR18	Basis of diagnosis
VAR19	Topography
VAR20	Morphology
VAR21	Behaviour
VAR22	Grade
VAR23	CIN grade **
VAR24	Multifocality **
VAR25	Laterality **
VAR26	Screen-detected **
Stage of disease at diagnosis	
VAR27	Pathological T
VAR28	Pathological N
VAR29	Pathological M
VAR30	Clinical T
VAR31	Clinical N
VAR32	Clinical M
VAR33	Site of metastasis **
VAR34	SEER Summary Stage 2000
VAR35	Condensed T
VAR36	Condensed N
VAR37	Condensed M
VAR38	FIGO stage
VAR39	Tumour size (mm)
VAR40	Sentinel lymph node biopsy **
VAR41	No. of lymph nodes examined
VAR42	No. of lymph nodes involved

Name	Short description
Diagnostic procedures **	
VAR43	Mammography
VAR44	Tissue diagnosis
VAR45	Papanicolaou (Pap) test
VAR46	HPV test
VAR47	Colposcopy
VAR48	Chest X-ray
VAR49	Abdominal ultrasound
VAR50	Scintigraphy
VAR51	Skeletal X-ray
VAR52	Computerised Tomography (CT) scan
VAR53	Trans-vaginal ultrasound
Biomarkers **	
VAR54	Oestrogen (Estrogen) receptors
VAR55	Progesterone receptors
VAR56	HER-2 based on IHC assay
VAR57	HER-2 based on FISH assay
VAR58	Ki-67 proliferation index
VAR59	CA-125
VAR60	*BRCA-1* or *BRCA-2*
Initial course of treatment	
VAR61	Emergency presentation **
VAR62	Cancer-directed surgery
VAR63	Day of first cancer-directed surgery
VAR64	Month of first cancer-directed surgery
VAR65	Year of first cancer-directed surgery
VAR66	Type of first cancer-directed surgery **
VAR67	Place of first cancer-directed surgery **
VAR68	Lymphadenectomy **
VAR69	Day of lymphadenectomy **
VAR70	Month of lymphadenectomy **
VAR71	Year of lymphadenectomy **
VAR72	Radiotherapy
VAR73	Day of first radiotherapy
VAR74	Month of first radiotherapy
VAR75	Year of first radiotherapy
VAR76	Total radiotherapy dose (Gy) **
VAR77	No. of radiotherapy fractions **
VAR78	Chemotherapy
VAR79	Day of first chemotherapy
VAR80	Month of first chemotherapy
VAR81	Year of first chemotherapy
VAR82	No. of chemotherapy cycles **
VAR83	Endocrine treatment **
VAR84	Year of first endocrine treatment
VAR85	Anti-HER-2 treatment **
VAR86	Day of first anti-HER-2 treatment
VAR87	Month of first anti-HER-2 treatment
VAR88	Year of first anti-HER-2 treatment
VAR89	Patient recruited to a clinical trial **

Socio-economic status **

VAR90 Education
VAR91 No. of persons living in household
VAR92 Home ownership
VAR93 Water supply
VAR94 Electricity supply
VAR95 Sanitation
VAR96 Socio-economic group
VAR97 Marital status
VAR98 Employment status
VAR99 Health insurance status

Family history **

VAR100 Breast or ovarian cancer in female 1st-degree relative
VAR101 No. of affected relatives

Lifestyle **

VAR102 Alcohol consumption
VAR103 Smoking habit
VAR104 Height (cm)
VAR105 Weight (kg)

Concurrent medical conditions **

VAR106 Cardiovascular conditions
VAR107 Diabetes
VAR108 Previous cancer
VAR109 Renal impairment
VAR110 Liver impairment
VAR111 History of HIV

* If your registry does not collect these data, please see paragraph 1.13
** New variable(s) in VENUSCANCER (not included in CONCORD-3)

2.2 Abbreviations

ASCO/CAP	American Society of Clinical Oncology and the College of American Pathologists
ASTRO	American Society for Radiation Oncology
BRCA1/2	BReast CAncer genes 1 and 2
CA-125	Cancer Antigen 125
CIN	Cervical intra-epithelial neoplasia
CT	Computed tomography
DCO	Death-certificate-only
ER	[O]estrogen Receptors
ESMO	European Society for Medical Oncology
FIGO	Fédération Internationale de Gynécologie et d'Obstétrique
HER-2	Human Epidermal growth factor Receptor 2
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
ICD-O	International Classification of Diseases for Oncology
IHC	Immunohistochemistry
ISH	*in situ* hybridization
Ki-67	Ki-67 proliferation index
LSHTM	London School of Hygiene and Tropical Medicine
MRI	Magnetic Resonance Imaging
NOS	Not Otherwise Specified
PgR	Progesterone Receptors
SEER	Surveillance, Epidemiology and End Results program
TNM	Tumour-Nodes-Metastasis
UICC	Union for International Cancer Control

3 DESCRIPTION OF VARIABLES

DEMOGRAPHICS

Country, age at diagnosis, sex, and race/ethnicity are key determinants of survival, and factors that contribute to differences in the provision of standard treatment.

Variable 1 **Country** **Core variable**

Numeric variable, four digits.

We will send you the code to be used for your country, to help you prepare your data. This 4-digit code must be included as variable 1 in **all tumour records**.

Variable 2 **Registry** **Core variable**

Numeric variable, one to three digits (range 1-950)

We will send you the code to be used for your registry, to help you prepare your data. The code for your registry must be included as variable 2 in **all tumour records**.

Together with the country code (variable 1), this variable will be used to link your data files with the relevant life tables during survival analysis.

Variable 3 **Person code** **Core variable**

Numeric variable, up to a maximum of 15 digits, ***or***
Alphanumeric variable, up to a maximum of 15 characters.

This is a unique code that can be used in your **cancer registry** to refer to each registered cancer patient.

The **person code** can be any unique string of characters, but ***not*** the person's name, national identity number, social security number or any similarly recognisable code.

The **person code** must be included in each tumour record, to enable you to check the record in the event that we identify possible errors during quality control. The same **person code** must be included in any other tumour records supplied for the same person.

Together with the **tumour code** (variable 4), this variable provides a unique identification of each tumour included in the study, for the purposes of quality control, but without compromising patient confidentiality.

Note:
If your registry does not routinely use a **person code**, you will need to create a unique code for each person to be included in your data files. The code will be used to identify patients with more than one index cancer. Also, if we identify errors or inconsistencies in a tumour record for a particular person during quality control, you will be able to use the person code to check each tumour record for that person in your registry.

If you have any doubts about the appropriate procedure to create a person code, please discuss this with us **before submitting your data**.

Variable 4 **Tumour code** **Core variable**

Numeric variable, up to a maximum of ten digits, ***or***
Alphanumeric variable, up to a maximum of ten characters.

This is the code used in your **cancer registry** to refer to each registered tumour.

Together with the **person code** (variable 3), this variable will enable persons with more than one **index cancer** to be identified.

For example, if you submit data for all patients diagnosed during 2015, a patient with an invasive primary cancer of the breast diagnosed in 2015, followed by a new invasive primary cancer of the cervix diagnosed in 2015, will be included in the survival analyses for each of those cancers.

Variable 5 **Sex** **Core variable**

Numeric variable, one digit.

Please do *not* select tumours for submission on the basis of sex: although we are interested in women's cancers, it will be of interest to monitor the proportion of breast and peritoneal cancers in men, as well as the proportion of records with unknown sex for each of the three index cancers.

Please do *not* exclude records from your data if the sex of the patient is unknown.

Code		Meaning
1	=	Male
2	=	Female
9	=	Sex is ambiguous, or sex was not known

Variable 6 **Region** **Optional variable**

Numeric variable, up to five digits.

It may be possible to examine patterns of care and to estimate survival for geographic areas within the territory of your registry. If your registry covers a province or state, survival analyses could be for smaller regions (e.g. counties) within your territory. Alternatively, areas may be classified as urban or rural.

You will need to include a suitable geographic code in **all tumour records**.

We give two examples below. Please discuss the codes you wish to use before submitting your data. You will need to tell us which region (county, etc.) corresponds to each code.

Code		Meaning
21001	=	Kentucky, Adair
21003	=	Kentucky, Allen
21...	=	Kentucky, ...
99999	=	Region not known, or you will not supply this variable

Code		Meaning
1	=	Urban
2	=	Rural
9	=	Region not known, or you will not supply this variable

Note:
We will need to construct life tables for each region (county, urban/rural areas, etc.) for which you wish to obtain separate survival estimates: see **Life tables (**CONCORD protocol **Annex 2)**.

Variable 7 **Race/ethnicity** **Optional variable**

Numeric variable, one or two digits.

It may be possible to estimate survival separately for each race/ethnicity within a population.

Cancer registries in some countries collect information on race and/or ethnicity (Australia, Israel, Malaysia, New Zealand, Singapore, UK, USA) or nationality (Kuwait). By contrast, most European registries **do not** record information on race or ethnicity. In some countries, it is illegal to do so.

Please use the code 99 for this variable in ***all tumour records***, if:

- you ***do not collect*** data on race or ethnicity
- you ***do collect*** data on race or ethnicity, but you decide not to provide this variable

If you would like us to examine patterns of care and to provide survival estimates by race or ethnicity, you will need to include a suitable code for race/ethnicity in **all tumour records**.

The categories for race/ethnicity will be different for each registry that supplies this variable.

You will need to tell us which race or ethnic group corresponds to each code. As an example, the codes usually used for race/ethnicity in the USA are:

Code		Meaning
1	=	White, Hispanic
2	=	White, Non-Hispanic
3	=	White, Hispanic status unknown
4	=	Black, Hispanic
5	=	Black, Non-Hispanic
6	=	Black, Hispanic status unknown
7	=	Asian or Pacific Islander, Hispanic
8	=	Asian or Pacific Islander, Non-Hispanic
9	=	Asian or Pacific Islander, Hispanic status unknown
10	=	American Indian/Alaska Native, Hispanic
11	=	American Indian/Alaska Native, Non-Hispanic
12	=	American Indian/Alaska Native, Hispanic status unknown
13	=	Other, unspecified or unknown race, Hispanic
14	=	Other, unspecified or unknown race, Non-Hispanic
15	=	Other, unspecified or unknown race, Hispanic status unknown
99	=	Unknown or missing, or variable not supplied

Note: If you would like us to provide survival analyses by race or ethnicity, we will need to construct life tables for each racial or ethnic group: see Life tables (CONCORD protocol Annex 2).

If we cannot construct robust life tables for each race or ethnic group, the scope for estimation of net survival by race or ethnic group may be limited.

Variables 8-10 Date of birth

A full and accurate date of ***birth*** (day, month and year) is important, because it is used to calculate the exact age at diagnosis. This is used to determine the age group (at diagnosis) into which patients are assigned for age-specific survival estimates. It is also used to calculate the exact age at death, and thus to select the appropriate background death rate from the life table in estimating net survival. It is also the basis for age-standardised survival.

Most registries record the full date of birth, but some registries face legal, ethical or regulatory problems in providing the full date of birth to external researchers. Some registries only record part of the date.

We have published a peer-reviewed article setting out why full dates are important, based on empirical evidence and sensitivity analyses with a large national data set.[7] The article shows the difficulties that arise in quality control when full dates cannot be obtained. It also shows the biases that arise in the estimation and comparison of survival. The article is accessible with your password on the CONCORD web-site.

The Cancer Survival Group at LSHTM maintains **statutory approval (**CONCORD protocol **Annex 4)** and **ethical approvals (**CONCORD protocol **Annex 5)** from the relevant bodies in the UK to receive and analyse individual tumour records with full dates of birth, diagnosis and death for the CONCORD programme.

If you cannot include full dates of birth, diagnosis and death, please discuss this with us before submitting your data. We may be able to help you obtain ethical approval, or to find an alternative solution.

Imputation of dates

Please do not impute the missing components of any dates. We will do this, if necessary, in the same way for all data sets. If missing components of some dates have *already* been imputed in your data, **we will ask you to tell us** the rules you use to impute the day and/or the month of any dates

If you routinely add a "flag" to tumour records to show when the day and/or the month of any date has been imputed, please discuss this with us **before submitting your data**.

Variable 8 Day of birth **Core variable**

Numeric variable, one or two digits.

1-31 = The day of birth
99 = The day of birth of this patient is not known

Variable 9 Month of birth **Core variable**

Numeric variable, one or two digits.

1-12 = The month of birth
99 = The month of birth of this patient is not known

Variable 10 Year of birth **Core variable**

Numeric variable, four digits.

YYYY = The year of birth
9999 = The year of birth of this patient is not known

Variables 11-13 Date of diagnosis

The date of diagnosis should be the date that is used by the registry for cancer incidence.

A full and accurate **date of diagnosis** (day, month, year) is important, because it is the starting point for the duration of survival. A few cancer registries only record the month and year of diagnosis. Other registries only began to record the *full* date of diagnosis at some point since 2000.

Variable 11 Day of diagnosis **Core variable**

Numeric variable, one or two digits.

1-31 = The day of the date of diagnosis
99 = The day of this date is not known

Note: please see comments above (page 8) about the imputation of dates.

Variable 12 Month of diagnosis **Core variable**

Numeric variable, one or two digits.

1-12 = The month of the date of diagnosis
99 = The month of this date is not known

Note: please see comments above (page 8) about the imputation of dates.

Variable 13 Year of diagnosis **Core variable**

Numeric variable, four digits.

YYYY = The year of diagnosis
Missing values are not allowed

FOLLOW-UP FOR VITAL STATUS

Variable 14 Last known vital status **Core variable**

Numeric variable, one digit.

This variable encodes the patient's **last known vital status**, to the extent that it is known to the cancer registry. The ***date*** of the patient's **last known vital status** is recorded in variables 15-17.

Code		**Meaning**
1	=	Alive
2	=	Dead
3	=	Lost to follow-up
9	=	Vital status is not known

Information about vital status is conventionally captured using either **'active'** or **'passive'** procedures, which we discuss below. Some registries use both.

'Active' follow-up
Active follow-up refers to the situation in which the registry actively seeks information about the vital status of each cancer patient on a regular basis, e.g. from the patient's doctor, or hospital, or even home visits.

If your registry uses this approach, you should use code "3" for patients whose vital status (alive or dead) could not be ascertained at the last vital status check: these patients are lost to follow-up. The last *date at which they were known to be alive* should be given in variables 15-17.

Code "3" should also be used for patients known to have emigrated, since they are also lost to follow-up. The *date of emigration* should be given in variables 15-17.

'Passive' follow-up
Passive follow-up refers to the situation in which the registry routinely receives information from one or more reliable sources on the vital status of *all registered patients*. These sources vary widely between registries, but may include social security or health insurance files, or a regional or national index of persons who have died. The information may be derived in various ways, such as by computer linkage with the registry database, manual scanning of the death index, or supply of details about all deaths in the registry's territory.

For international survival comparisons, the key features of passive follow-up are that:
- The registry uses this approach for updating its own data for survival analysis, and
- The registry's procedures reliably identify ***all deaths of registered cancer patients, not just the deaths for which cancer is mentioned on the death certificate***, and
- The registry can reliably assume that registered cancer patients are alive, unless information about a patient's death has been received from one or more of these sources.

If your registry uses passive follow-up, patients who are *not known to be dead* would normally be assumed to be alive on the date of the most recent linkage between the registry and a death index or other vital status records. The vital status of those patients should be coded as "1" (alive).

If some patients cannot be traced by any passive follow-up procedure, their vital status may remain undetermined. It should then be coded as "9" (unknown).

When you submit your data, we will ask you to tell us *the last date when you performed linkage* or other follow-up to determine the vital status of your patients (**the "freeze" date**).

Variables 15-17 Date of last known vital status

This is the most recent date for which the patient's **last known vital status** (variable 14) was available.

The **date of last known vital status** should be coded as:

- the date of death - if the patient is dead (variable 14 should be coded as "2")
- missing (99,99,9999) - if the patient is known to be dead, but the date of death is not known
- the date of emigration - if the patient has emigrated (variable 14 is coded as "3")
- the date of loss to follow-up - if the patient has been lost to follow-up (variable 14 is coded as "3")

If the patient is considered to be alive, but not as emigrated or lost to follow-up, the **date of last known vital status** should be one of the following:

- 31 December of the last year for which follow-up of all patients is believed to be complete, ***or***
- the date on which the registry last checked that patient's vital status, e.g. contact with the patient's doctor or a home visit (registries that perform active follow-up), ***or***
- the date of the most recent linkage with a death index ("freeze" date; registries that perform passive follow-up)

Variable 15 Day of last known vital status **Core variable**

Numeric variable, one or two digits.

1-31 = The day of the date of last known vital status
99 = The day of this date is not known

Note: please see comments above (page 8) about the imputation of dates.

Variable 16 Month of last known vital status **Core variable**

Numeric variable, one or two digits.

1-12 = The month of the date of last known vital status
99 = The month of this date is not known

Note: please see comments above (page 8) about the imputation of dates.

Variable 17 Year of last known vital status **Core variable**

Numeric variable, four digits.

YYYY = The year of the date of last known vital status
9999 = The year of this date is not known

TUMOUR DETAILS

Variable 18 Basis of diagnosis Core variable

Numeric variable, one digit.

This variable indicates the degree of certainty with which a diagnosis of cancer has been established.

The distinction between a diagnosis based on pathological examination of cells or tissue (microscopically verified) and a diagnosis based on other investigations, such as clinical examination or imaging, is important for both cancer incidence and survival. However, the categories for the basis of diagnosis required for survival analysis differ slightly from the categories conventionally used for incidence ***(see note below)***.

Code Meaning

Not microscopically verified

1 = Clinical diagnosis ***only***

2 = Clinical investigation: endoscopy (without biopsy) or imaging (e.g. X-ray, ultrasound, scintigraphy, computed tomography (CT) or magnetic resonance imaging (MRI))

3 = Clinical diagnosis, not otherwise specified (i.e. not known if code "1" or "2" applies)

Microscopically verified

4 = Cytologically confirmed (e.g. smear; aspirate of bone marrow or ascites)

5 = Histologically confirmed (tissue diagnosis from biopsy or surgical specimen)

6 = Microscopically verified, not otherwise specified (i.e. not known if code "4" or "5" applies)

Evidence of cancer does not include the date of diagnosis

7 = Death-certificate-only registration (DCO) ***(see note below)***

8 = Autopsy only - malignancy detected only at autopsy ***(see note below)***

No information

9 = Unknown

Note:
Cancers registered solely on the basis of a death certificate (code 7) or detected only at autopsy (code 8) can usually be **included in incidence statistics** for the year in which they are registered.

For DCO and autopsy-detected cancers, however, the true date of diagnosis – and thus the duration of survival – are unknown. Therefore, they are normally ***excluded from survival analyses***. Some cancer registries do not register DCOs or autopsy-detected cancers.

If your cancer registry ***did*** register DCO or autopsy-only cancers during the calendar period covered by your data submission, **you should include those cancers in the data you submit for this study**, to enable comparative quality control.

Variable 19 Topography **Core variable**

Alphanumeric variable, four characters.

Please select records for submission with the tumour site (topography), using the following codes in the International Classification of Diseases for Oncology (ICD-O-3[2]; topography codes are the same in the first revision of ICD-O-3[6]).

Please provide the full 4-character ICD-O-3 topography code ***without the decimal point (".")***. Thus, for example, breast cancer of upper-inner quadrant will be C502. With this modification, the anatomic site of the **index cancers** will be coded as:

Breast: C500-C506; C508-C509
Note: excludes skin of breast, C445

Cervix uteri: C530-C531; C538-C539

Ovary: C480-C482; C488; C569; C570-C574; C577-C579

Note: includes peritoneum and retroperitoneum (C480-C482), where high-grade serous cancers originating from the fallopian tube (C570) are often detected
Note: includes overlapping lesion of retroperitoneum and peritoneum (C488)
Note: includes other and unspecified female genital organs (C577-C579)

Variable 20 Morphology **Core variable**

Numeric variable, four digits.

Tumour morphology should be coded to the third edition of the International Classification of Diseases for Oncology (ICD-O-3),[2] or the first revision of ICD-O-3.[6]

For microscopically confirmed tumours, the range of morphology codes in ICD-O-3 (as updated by the first revision) is now **8000-9992**.

For tumours *without* microscopic verification, you should use:

9999 This is ***not*** a valid ICD-O-3 code. However, we ask you to use it here to mean that for tumours without microscopic verification, you do not have morphology data.

Variable 21 Behaviour **Core variable**

Numeric variable, one digit.

Please do *not* select tumours for submission on the basis of tumour behaviour. If your registry collects data on tumours that are benign (behaviour code 0), of uncertain behaviour (1) or *in situ* (2), such as *in situ* carcinoma of the cervix, ***please include all these records*** in your data files.

We will report the **distribution of tumour behaviour** for each cancer. This will enable comparison of the intensity of diagnostic activity between participating countries or regions, e.g. the proportion of women with breast or cervical cancer who were registered with *in situ* carcinoma.

Survival analyses will ***only*** include invasive, primary, malignant neoplasms (behaviour code 3). We may include *in situ* tumours in patterns of care analyses.

Tumour behaviour should be coded to the third edition of the International Classification of Diseases for Oncology (ICD-O-3).[2] The coding of tumour behaviour has been the same in all editions of ICD-O.

Code		Meaning
0	=	Benign
1	=	Uncertain whether benign or malignant
2	=	Carcinoma *in situ*
3	=	Malignant, primary site
6	=	Malignant, metastatic site ***(see note below)***
9	=	Malignant, uncertain whether primary or metastatic site ***(see note below)***

Note:
Behaviour codes 6 and 9 are included in ICD-O-3, but they are not usually used by cancer registries.[2] We show them here only for completeness.

If your registry ***does*** use behaviour codes 6 and 9, please include them in your data files; do not recode them before data submission. Instead, ***please provide us with a description*** of how these codes have been used in your data.

Variable 22 Grade **Core variable**

Numeric variable, one digit.

Histological grade (or differentiation) is a major prognostic factor for several cancers. It is only applicable to malignant tumours. It should reflect the highest differentiation code recorded in the diagnostic statement or pathological report, i.e. the one with the most adverse prognostic significance (Rule G in ICD-O-3[6]).

If you do not collect data on histological grade or differentiation, please use the code 9 for this variable in ***all tumour records***.

Code		Meaning	
1	=	Grade I	Well differentiated, ***or*** Differentiated, NOS
2	=	Grade II	Moderately differentiated, ***or*** Moderately well differentiated, ***or*** Intermediate differentiation
3	=	Grade III	Poorly differentiated
4	=	Grade IV	Undifferentiated, ***or*** Anaplastic
8	=	No pathology was performed	
9	=	Grade or differentiation not determined, ***or*** Not stated	

Variable 23 CIN grade **Core variable**

Numeric variable, one digit.

Cervical intraepithelial neoplasia (CIN) is a precancerous condition in which abnormal cells grow on the surface of the cervix. If not treated, these abnormal cells may become invasive and spread to nearby normal tissue. Experts in this field argue that CIN III should be considered as comparable with carcinoma in situ (behaviour code /2), even if severe dysplasia is not mentioned. Pathologists who do not accept this equivalence may assign the behaviour code /1 (uncertain whether benign or malignant).[6]

If you do not collect data on CIN grade, please use code 9 in ***all tumour records*** referring to cervical tumour.

For women diagnosed with breast or ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	CIN 1: dysplasia that involves about one-third of the thickness of the epithelium
2	=	CIN 2: abnormal changes in about one-third to two-thirds of the epithelial layer
3	=	CIN 3 (the most severe form) describes a condition that affects more than two-thirds of the epithelium
8	=	No pathology was performed
9	=	CIN grade not determined, ***or*** Not stated, ***or*** Not applicable (***breast, ovary***)

Variable 24 **Multifocality** **Optional variable**

Numeric variable, one digit.

Several tumours that are apparently not in continuity with other tumours originating in the same primary site or tissue[6] may be observed. In the breast, these multifocal (or multicentric) lesions may arise in the same quadrant or different quadrants of the same breast.[8,9]

Clinically, it is important to determine the nature of multiple lesions, because they may be indicative of a familial cancer,[10] and they may have a different prognosis from that of unifocal lesions.[11]

Code		Meaning
1	=	***Presence*** of multifocality/multicentricity
2	=	***Absence*** of multifocality/multicentricity
3	=	No pathology was performed
9	=	Multifocality/multicentricity not determined, ***or*** Not stated

Variable 25 **Laterality** **Optional variable**

Numeric variable, one digit.

Breast and ovary are paired (bilateral) organs. Laterality describes the side of the body in which a tumour of a paired organ originated.

Cancer registries may use different rules to collect information on laterality.[6,12]

For women diagnosed with cervical cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Unilateral tumour, right side
2	=	Unilateral tumour, left side
3	=	Bilateral tumour
9	=	Laterality not determined, ***or*** Not stated, ***or*** Not applicable (***cervix***)

Variable 26 **Screen-detected** **Optional variable**

Numeric variable, one digit.

International guidelines recommend that national, population-based screening programmes should be implemented for breast and cervical cancers, and that most women in the target age range have access to them.

For women diagnosed with ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Yes, the tumour ***was screen-detected***
2	=	No, the tumour ***was not screen-detected***
3	=	Screening programme was not available
9	=	Not stated if the tumour was screen-detected, ***or*** Not applicable (***ovary***)

STAGE OF DISEASE AT DIAGNOSIS

"Tumour stage" describes how far the cancer has spread at the time of diagnosis. It is a key determinant of patient management and survival.

Knowledge of the stage at diagnosis is important to evaluate adherence to internationally agreed clinical guidelines[13-16] and for the interpretation of international survival comparisons.[17-20]

Many coding schemes are used to categorise tumour stage in cancer registries around the world.

Tumour-Nodes-Metastasis (TNM) is the preferred option. However, we will try to obtain data on **at least one widely used categorisation of stage** at diagnosis for each tumour:

***** Data on TNM stage must be coded according to the **7th edition.**[21]

If possible, please supply data for ***at least one*** of the stage classifications in the left-hand column.

For variables on stage at diagnosis that you do not collect, please use the relevant code for missing values (see page 46) ***in all tumour records.*** For example, if you submit data for TNM stage, but you do not collect SEER Summary Stage 2000, you should code "9" for SEER Summary Stage 2000 in all tumour records.

Variables 27-32 TNM stage

The Tumour-Nodes-Metastasis (TNM) classification of stage at diagnosis uses a combination of clinical and pathological evidence, like SEER Summary Stage 2000.

Please use the **7th edition**[21] **of TNM** (UICC)**.** This is identical to the classification published by the American Joint Committee on Cancer in 2010.[22] If your stage data are coded to earlier editions than TNM-7, please contact us **before submitting your data.**

The three components of TNM are tumour size or local extension (T); the status of regional lymph nodes, i.e. the extent of lymph node invasion by tumour (N), and whether there is metastasis (spread of disease to an organ or organs distant from the organ of origin) (M).

TNM stage may be based on pathological evidence ("p") or clinical evidence ("c").

Variables 27-29 Pathological TNM stage

Variable 27 Pathological T — Core variable

Numeric variable, one digit.

This variable encodes information on the physical size or extent of the tumour.

For all three index cancers, up to 4 subcategories (a, b, c, d) exist for each of the stage categories pT1, pT2, pT3 and pT4. These subcategories should be coded in the same way as the parent category: for example, pT1a should be coded to "1", in the same way as pT1.

For **cervical cancer**, the additional sub-categories pT1a1, pT1a2, pT1b1 and pT1b2 should be coded to "1", in the same way as pT1. Similarly, the additional sub-categories pT2a1 and pT2a2 should be coded to "2", in the same way as pT2.

Code **Meaning**

0 = pT0 – no histological evidence of primary tumour

1 = pT1 – the content of this category varies with the cancer[21]
Use this code also for sub-categories pT1a, pT1b and pT1c
*Use this code also for sub-categories pT1a1, pT1a2, pT1b1 and pT1b2 (**cervix**)*

2 = pT2 – the content of this category varies with the cancer[21]
Use this code also for sub-categories pT2a, pT2b and pT2c
*Use this code also for sub-categories pT2a1 and pT2a2 (**cervix**)*

3 = pT3 – the content of this category varies with the cancer[21]
Use this code also for sub-categories pT3a, pT3b and pT3c

4 = pT4 – the content of this category varies with the cancer[21]
Use this code also for sub-categories pT4a, pT4b, pT4c and pT4d

8 = *is – in situ* carcinoma

9 = pTX – unknown: the primary tumour cannot be assessed histologically

Variable 28 Pathological N — Core variable

Numeric variable, one digit.

This variable encodes the extent of involvement of regional lymph nodes with tumour.

For cancers of the breast, 3 subcategories of pN1, pN2 and pN3 also exist (a, b and c). These should be coded in the same way as the parent category: for example, pN2b should be coded to "2", in the same way as pN2.

Code **Meaning**

0 = pN0 – no regional lymph nodes involved with tumour, histologically

1 = pN1 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories pN1a, pN1b and pN1c (**breast**)*
*Use this code also for pN1mi (micrometastasis) (**breast**)*

2 = pN2 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories pN2a and pN2b (**breast**)*

3 = pN3 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories pN3a, pN3b and pN3c (**breast**)*

9 = pNX – unknown: the regional lymph nodes cannot be assessed histologically

Variable 29 Pathological M **Core variable**

Numeric variable, one digit.

This variable encodes information on the presence or absence of distant metastases.

Code		Meaning
1	=	pM1 – Distant metastases have been microscopically confirmed
9	=	Unknown ***Note***: this is not a valid TNM code (***see below***): we use it here only to show that no data were available on pathological M status for this tumour.

The code "MX" was used in earlier editions of TNM to indicate that the metastatic status of the tumour was unknown. However, clinical assessment of metastasis can be based on physical examination alone, so pMX is no longer considered an appropriate code.

The codes pM0 and pMX are *not valid* in the 7th edition of TNM.[21]

Variables 30-32 Clinical TNM stage

Clinical data on tumour stage may be available from clinical examination, or from imaging of the tumour (X-ray, computed tomography [CT], magnetic resonance imaging [MRI], ultrasound, etc.). If no surgical or invasive diagnostic procedure has been performed, clinical data may be the only available data on tumour stage.

Variable 30 Clinical T **Core variable**

Numeric variable, one digit.

This variable encodes information on the physical size or extent of the tumour.

For all three index cancers, up to 4 sub-categories (a, b, c, d) exist for each of the stage categories cT1, cT2, cT3 and cT4. These sub-categories should be coded in the same way as the parent category: for example, cT1a should be coded to "1", in the same way as cT1.

For **cervical cancer**, the additional sub-categories cT1a1, cT1a2, cT1b1 and cT1b2 should all be coded to "1", in the same way as cT1. Similarly, the additional sub-categories cT2a1 and cT2a2 should both be coded to "2", in the same way as cT2.

Code		Meaning
0	=	cT0 – no evidence of primary tumour
1	=	cT1 – the content of this category varies with the cancer[21] *Use this code also for sub-categories cT1a, cT1b and cT1c* *Use this code also for sub-categories cT1a1, cT1a2, cT1b1 and cT1b2* (***cervix***)
2	=	cT2 – the content of this category varies with the cancer[21] *Use this code also for sub-categories cT2a, cT2b and cT2c* *Use this code also for sub-categories cT2a1 and cT2a2 (**cervix**)*
3	=	cT3 – the content of this category varies with the cancer[21] *Use this code also for sub-categories cT3a, cT3b and cT3c*
4	=	cT4 – the content of this category varies with the cancer[21] *Use this code also for sub-categories cT4a, cT4b, cT4c and cT4d*
8	=	*is* – in situ carcinoma
9	=	cTX – unknown: the primary tumour cannot be assessed

Variable 31 Clinical N **Core variable**

Numeric variable, one digit.

This variable encodes information on the involvement of regional lymph nodes with tumour.

For cancers of the breast, three subcategories of cN1, cN2 and cN3 also exist (a, b and c). These should be coded in the same way as the parent category: for example, cN2b should be coded to "2", in the same way as cN2.

Code Meaning

0 = cN0 – no regional lymph nodes involved with tumour, histologically[21]

1 = cN1 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories cN1a, cN1b and cN1c (**breast**)*

2 = cN2 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories cN2a and cN2b (**breast**)*

3 = cN3 – the content of this category varies with the cancer[21]
*Use this code also for sub-categories cN3a, cN3b and cN3c (**breast**)*

9 = cNX – unknown: the regional lymph nodes cannot be assessed

Variable 32 Clinical M **Core variable**

Numeric variable, one digit.

This variable encodes information on the presence or absence of distant metastases.

Code Meaning

0 = cM0 – No metastases ***(see note 1 below)***

1 = cM1 – Metastases ***(see note 2 below)***

9 = This is ***not*** a valid code in the 7th edition of TNM. We ask you to use it here only if you are **not providing data on TNM stage**.

Note 1:
If the clinician ***does not record the presence of metastases***, it is assumed under TNM-7[21] that no metastases are present (cM0). You should code these cases to "0".

Note 2:
The code "MX" was used in earlier editions of TNM to indicate that the metastatic status of the tumour was unknown. However, assessment of metastasis can be based on clinical examination or imaging alone, so cMX is no longer considered an appropriate code.

The code cMX is not valid in the 7th[21] edition of TNM.

Variable 33 Site of metastasis **Optional variable**

Numeric variable, one digit.

For breast cancer only. This variable encodes information on the site of distant metastases, derived from the 7th edition of TNM.[21] For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records.***

If more than one metastatic site is involved, select the lowest code. For example, if metastases are detected in the brain and the lung (visceral), use the code for brain (1).

Code		Meaning
1	=	Brain
2	=	Visceral *
3	=	Non-visceral **
9	=	Site of metastasis is unknown, ***or*** Not stated, ***or*** Not applicable (no metastasis), ***or*** Not applicable (***cervix, ovary***)

***** **Visceral** – lung, pleura, liver, peritoneum, adrenal, other internal organs or bone marrow
****** **Non-visceral** – bone, lymph node or skin

Variable 34 SEER Summary Stage 2000 **Core variable**

Numeric variable, one digit.

SEER Summary Stage 2000 is a simple categorisation of stage, developed by the US National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) program. The North American Association of Central Cancer Registries uses SEER Summary Stage 2000[23] (http://seer.cancer.gov/tools/ssm/). We borrow text from the introduction here. It has been in use in the US and Canada since 1 January 2001.

"Summary stage" is the most basic way of categorising how far a cancer has spread from its point of origin. Summary staging uses all the information available in the medical record. It is a **combination of the most precise clinical and pathological evidence for the extent of disease**. Many population-based cancer registries report summary stage, because the staging categories are sufficiently broad to enable measurement of progress in cancer control.

Regional spread of disease is divided into several categories, according to the method of spread of the cancer:

Code		Meaning
0	=	*In situ*
1	=	Localised only
2	=	Regional spread by direct extension ***only***
3	=	Regional lymph nodes involved ***only***
4	=	Regional spread by ***both*** direct extension ***and*** lymph node involvement
5	=	Regional, not otherwise specified – use this code if there is regional spread of the cancer, but ***the route of spread is not known***
7	=	Distant site(s) or lymph node(s) are involved
9	=	Unknown if there is extension or metastasis (unstaged, unknown or unspecified), ***or*** This is a ***death-certificate-only*** case, ***or*** This is an ***autopsy-only*** case

Variables 35-37 Condensed TNM stage

The condensed TNM scheme for recording tumour stage was developed by the European Network of Cancer Registries[24] for tumour records in which the individual values of T or N or M are not explicitly recorded. Condensed TNM is based on the TNM 6th edition.[25]

Please use code 9 for variables 35-37 in ***all tumour records if***:

- your registry ***does not use*** condensed TNM
- your registry ***does use*** condensed TNM, but you decide not to provide these variables

Variable 35 Condensed T — Core variable

Numeric variable, one digit.

Code **Meaning**

1 = L - Localised disease
Use this code for:
T1 cancers of the ovary
T1 and T2 cancers of the cervix
T1, T2 and T3 cancers of the breast

2 = A - Advanced disease
Use this code for:
T2 and T3 cancers of the ovary
T3 and T4 cancers of cervix
T4 cancers of the breast

9 = X – Cannot be assessed: no information on tumour size category

Variable 36 Condensed N — Core variable

Numeric variable, one digit.

Code **Meaning**

1 = N0 – No regional lymph node invasion by tumour
2 = N+ – Regional lymph nodes invaded by tumour
9 = NX – Cannot be assessed: no information on nodal status

Variable 37 Condensed M — Core variable

Numeric variable, one digit.

This code is based on the best available information - clinical, instrumental or pathological. Clinical signs and findings may be sufficient to justify classifying a tumour as having metastasised (M+), even without pathological confirmation of metastatic deposits.

Code **Meaning**

1 = M0 – No distant metastasis
2 = M+ – Distant metastasis present
9 = MX – Cannot be assessed: no information on whether metastases are present

Variable 38 FIGO stage **Core variable**

Numeric variable, one digit.

FIGO stage[26] is a specialised classification of tumour stage for cervical and ovarian cancers.

FIGO stage should ***only*** be reported if TNM stage data (variables 27-32) are not available. The TNM classification is preferable, because it is more detailed.

FIGO stage provides five broad categories, which ***differ between cervix and ovary (see below):***

Stage 0	Carcinoma *in situ* (common in cervical cancer)
Stage I	Confined to the organ of origin
Stage II	Invasion of surrounding organs or tissue
Stage III	Spread to distant nodes or tissue within the pelvis
Stage IV	Distant metastasis(es)

For women diagnosed with breast cancer, please use code 9 in ***all tumour records.***

Cervix (C530-C531; C538-C539)

Code Meaning

0 = FIGO Stage 0 – Carcinoma *in situ*

1 = FIGO Stage I – Tumour confined to cervix (extension to corpus uteri should be disregarded)
Use this code also for sub-categories IA, IA1, IA2, IB, IB1 and IB2

2 = FIGO Stage II – Tumour invades beyond uterus but not to pelvic wall or lower third of vagina
Use this code also for sub-categories IIA, IIA1, IIA2 and IIB

3 = FIGO Stage III – Tumour extends to pelvic wall or lower third of vagina, or causes hydronephrosis
Use this code also for sub-categories IIIA and IIIB

4 = FIGO Stage IVA – Tumour invades mucosa of the bladder or rectum, or extends beyond true pelvis

5 = FIGO Stage IVB – Distant metastasis

9 = FIGO Stage unknown, ***or***
Not applicable (***breast***)

Ovary (C480-C482; C488; C569; C570-C574; C577-C579)

Code Meaning

1 = FIGO Stage I – Tumour limited to one or both ovaries

2 = FIGO Stage II – Tumour limited to one or both ovaries or fallopian tubes, with pelvic extension
Use this code also for sub-categories IIA, IIB and IIC

3 = FIGO Stage III – Tumour involves one or both ovaries with microscopically confirmed peritoneal metastasis outside the pelvis, and/or regional lymph node metastasis
Use this code also for sub-categories IIIA, IIIB and IIIC

4 = FIGO Stage IV – Distant metastasis outside the peritoneal cavity

9 = FIGO Stage unknown, ***or***
Not applicable (***breast***)

Variable 39 **Tumour size (mm)** **Core variable**

Numeric variable, from one to three digits.

Tumour size (maximum tumour diameter) should be reported in **millimetres**, as an integer.

For **breast cancer**, tumour size should be based on histological examination, if available.

It is difficult to be prescriptive about the maximum physical dimensions of a tumour. We will accept values in the range 1 - 300mm. For example, a tumour with a maximum diameter of 35mm (3.5cm) would be coded as "35". Zero is ***not*** a valid tumour dimension. Use code 999 if tumour size is not available.

Code		Meaning
1-300	=	Maximum tumour diameter, **in millimetres**
999	=	Maximum tumour diameter is unknown

Variable 40 **Sentinel lymph node biopsy** **Core variable**

Numeric variable, one digit.

Sentinel lymph node biopsy for breast and cervical cancers consists in finding and removing the first lymph node (or nodes) to which the tumour is likely to spread (the sentinel lymph nodes). [13-16]

For women diagnosed with ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Sentinel lymph node biopsy ***was performed***
2	=	Sentinel lymph node biopsy ***was not performed***
9	=	Not known if sentinel lymph node biopsy was performed, ***or*** Not applicable (***ovary***)

Variable 41 **No. of lymph nodes examined** **Core variable**

Numeric variable, one or two digits.

Report the exact ***number of lymph nodes examined***, as recorded in the pathological report: valid range 0-98.

Code		Meaning
0-98	=	Number of lymph nodes **examined**
99	=	The number of lymph nodes examined is unknown, ***or*** No pathological examination was done

Variable 42 **No. of lymph nodes involved** **Core variable**

Numeric variable, one or two digits.

Report the exact ***number of involved lymph nodes (lymph nodes containing tumour cells)***, as recorded in the pathological report: valid range 0-98.

For **breast and cervical cancers**, this number should also include any sentinel lymph nodes that are involved.

Code		Meaning
0-98	=	Number of **involved** lymph nodes
99	=	The number of **involved** lymph nodes is unknown, ***or*** No pathological examination was done

DIAGNOSTIC PROCEDURES

Provision of data on diagnostic procedures is *optional*.

For the index cancers, diagnosis is based on clinical examination, imaging and pathology.

The availability and intensity of diagnostic procedures contributes to the interpretation of international differences in the distribution of stage at diagnosis and stage-specific survival.[27]

Please use the codes for missing values (see page 46) for ***all the variables*** on diagnostic procedures (variables 43-53) ***in all tumour records if***:

- you ***do not collect*** data on diagnostic procedures
- you ***do collect*** data on diagnostic procedures, but you decide not to provide these variables

Variable 43 Mammography — Optional variable

Numeric variable, one digit.

This variable is specific for breast cancer. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Mammography ***was performed, usual type***
2	=	Mammography ***was performed, 3D or modern type***
3	=	Mammography ***was performed, unknown type***
4	=	Mammography ***was not performed***
9	=	Not known if mammography was performed, ***or*** Not applicable (***cervix, ovary***)

Variable 44 Tissue diagnosis — Optional variable

Numeric variable, one digit.

This variable is specific for breast and cervical cancers. For women diagnosed with ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Tissue diagnosis ***was performed, biopsy***
2	=	Tissue diagnosis ***was performed, needle aspiration (breast)***
3	=	Tissue diagnosis was not performed
9	=	Not known if tissue diagnosis was performed, ***or*** Not applicable (***ovary***)

Variable 45 **Papanicolaou (Pap) test** **Optional variable**

Numeric variable, one digit.

For cervical cancer only, whether a Pap test was performed during the diagnostic investigation. For women diagnosed with breast or ovarian cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	Pap test ***was performed***
2	=	Pap test ***was not performed***
9	=	Not known if Pap test was performed, ***or*** Not applicable (***breast, ovary***)

Variable 46 **HPV test** **Optional variable**

Numeric variable, one digit.

For cervical cancer only, whether the presence of HPV in the cervical smear was tested during the diagnostic investigation. For women diagnosed with breast or ovarian cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	HPV test ***was performed***
2	=	HPV test ***was not performed***
9	=	Not known if HPV test was performed, ***or*** Not applicable (***breast, ovary***)

Variable 47 **Colposcopy** **Optional variable**

Numeric variable, one digit.

For cervical cancer only. For women diagnosed with breast or ovarian cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	Colposcopy ***was performed***
2	=	Colposcopy ***was not performed***
9	=	Not known if colposcopy was performed, ***or*** Not applicable (***breast, ovary***)

Variable 48 **Chest X-ray** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Chest X-ray ***was performed***
2	=	Chest X-ray ***was not performed***
9	=	Not known if chest X-ray was performed

Variable 49 **Abdominal ultrasound** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Abdominal ultrasound ***was performed***
		Note: ***includes*** *ultrasound of the liver*
2	=	Abdominal ultrasound ***was not performed***
9	=	Not known if abdominal ultrasound was performed

Variable 50 **Scintigraphy** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Scintigraphy ***was performed***
2	=	Scintigraphy ***was not performed***
9	=	Not known if scintigraphy was performed

Variable 51 **Skeletal X-ray** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Skeletal X-ray ***was performed***
2	=	Skeletal X-ray ***was not performed***
9	=	Not known if skeletal X-ray was performed

Variable 52 **Computerised Tomography (CT) scan** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	CT scan ***was performed***
		Note: includes *Positron Emission Tomography (PET)*
2	=	CT scan ***was not performed***
9	=	Not known if CT scan was performed

Variable 53 **Trans-vaginal ultrasound** **Optional variable**

Numeric variable, one digit.

For cervical and ovarian cancer only. For women diagnosed with breast cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	Trans-vaginal ultrasound ***was performed***
2	=	Trans-vaginal ultrasound ***was not performed***
9	=	Not known if trans-vaginal ultrasound was performed, ***or***
		Not applicable (***breast***)

BIOMARKERS

Besides the classical clinico-pathological features (tumour size, morphology, grade and stage at diagnosis), various biomarkers influence the choice of treatment and the prognosis for **breast and ovarian** cancers.

Oestrogen (Estrogen) receptors (ER) and **progesterone** receptors (PgR), human epidermal growth factor receptor 2 (**HER-2**), and the **Ki-67** proliferation index are established biomarkers for ***breast cancer***.[13,14]

Mutations in **BReast CAncer genes one (*BRCA-1*) and two (*BRCA-2*)** increase the risk of breast and ovarian cancers, among other cancers.[13-16] As effective novel treatments for BRCA-related cancers have emerged, testing for BRCA mutations is increasingly widespread.[28,29]

CA-125 is the only tumour marker commonly assessed in the diagnosis and management of ***ovarian cancer***.

Variable 54 Oestrogen (Estrogen) receptors Core variable

Numeric variable, one digit.

The presence of oestrogen (estrogen) receptors (ER+) serves as a prognostic marker for responsiveness to endocrine treatment. Because the response to hormonal treatment in breast cancer is associated with ER positivity, assessment of receptor expression profile allows clinicians to predict breast cancer response to hormonal treatment.[13,14] Cut-offs for positivity may differ world-wide.

For breast cancer only. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records***.

Please code the results of this test **as defined in your territory**. According to the 2010 guidelines of the American Society of Clinical Oncology and the College of American Pathologists (ASCO/CAP),[30] possible ER test results are:

Code		Meaning
1	=	ER ***was tested, negative***
2	=	ER ***was tested, positive***
3	=	ER ***was not tested***
9	=	Not known if ER was tested, or Result not stated, or Not applicable (***cervix, ovary***)

Variable 55 Progesterone receptors Core variable

Numeric variable, one digit.

The presence of progesterone receptors (PgR+) also serves as a prognostic marker for the responsiveness of breast cancer to endocrine treatment.[13,14] Cut-offs for positivity may vary world-wide.

For breast cancer only. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records***.

Please code the results of this test **as defined in your territory**. According to the ASCO/CAP guidelines for 2010,[30] possible PgR test results are:

Code		Meaning
1	=	PgR was ***tested, negative***
2	=	PgR was ***tested, positive***
3	=	PgR was ***not tested***
9	=	Not known if PgR was tested, ***or*** Result not stated, ***or*** Not applicable (***cervix, ovary***)

Variable 56 **HER-2 based on immunohistochemistry (IHC) assay** **Core variable**

Numeric variable, one digit.

For breast cancer only. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records***.

We will refer to the ASCO/CAP 2013 guidelines[31] as the basis to collect data on HER-2 status. The guidelines and recommendations were updated in 2018. The data call for VENUSCANCER is being launched in 2019, but the new guidelines were not in force during the period covered by this study.

HER-2 status is tested by using either immunohistochemistry (IHC) or *in situ* hybridization, typically fluorescent *in situ* hybridisation (FISH).

According to the 2013 ASCO/CAP guidelines, IHC results can be negative (score 0 or 1+), equivocal (2+) or positive (3+).

On this basis, HER-2 test results based on IHC should be coded as follows:

Code		Meaning
1	=	HER-2 ***was tested, negative*** *Immunohistochemistry score 0 or 1+*
2	=	HER-2 ***was tested, equivocal*** *Immunohistochemistry score 2+*
3	=	HER-2 ***was tested, positive*** *Immunohistochemistry score 3+*
4	=	HER-2 ***was not tested***
9	=	Not known if HER-2 was tested, ***or*** Result not stated, ***or*** Not applicable (***cervix, ovary***)

Variable 57 **HER-2 based on immunofluorescence (FISH) assay** **Core variable**

Numeric variable, one digit.

For breast cancer only. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records***.

When immunohistochemistry (variable 56) is equivocal, fluorescent *in situ* hybridisation (FISH) or another second-level test is recommended.

In clinical records, results of FISH are usually reported as “positive”, “equivocal” or “negative”, determined on a single or dual probe. For more details, see ASCO/CAP guidelines 2013.[31]

Code		Meaning
1	=	Test was ***performed, negative***
2	=	Test was ***performed, equivocal***
3	=	Test was ***performed, positive***
4	=	Test was ***not performed***
9	=	Not known if test was performed, ***or*** Not stated, ***or*** Not applicable (***cervix, ovary***)

Variable 58 **Ki-67 proliferation index** **Optional variable**

Numeric variable, from one to three digits.

Ki-67 is a nuclear protein associated with cellular proliferation. For example, Ki-67 has recently been recommended for identification of luminal sub-types (A and B) of breast cancer. Cut-offs for positivity are still under debate, so Ki-67 scores should be coded according to the laboratory values in your territory.[13,32] We will collect the proportion of malignant cells staining positive for Ki-67.

Code		Meaning
0-100	=	Proportion of malignant cells staining positive for Ki-67
888	=	Ki-67 evaluation ***was not done***
999	=	Proportion of the malignant cells staining positive for Ki-67 is unknown, ***or*** Not stated

Variable 59 **CA-125** **Optional variable**

Numeric variable, one digit.

Measurement of serum CA-125 is still widely used to aid ovarian cancer diagnosis. It is also used to assess the response to chemotherapy and to diagnose possible recurrence.[15,16]

For women diagnosed with breast or cervical cancer, please use code 9 for this variable in ***all tumour records***.

Code		Meaning
1	=	CA-125 measurement ***was performed***
2	=	CA-125 measurement ***was not performed***
9	=	Not known if CA-125 measurement was performed, ***or*** Not applicable (***breast, cervix***)

Variable 60 ***BRCA-1* or *BRCA-2*** **Optional variable**

Numeric variable, one digit.

For breast and ovarian cancer only. For women diagnosed with cervical cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	*BRCA-1*/*BRCA-2* test ***was performed, mutation not present***
2	=	*BRCA-1*/*BRCA-2* test ***was performed, mutation present***
3	=	*BRCA-1*/*BRCA-2* test ***was not performed***
9	=	Not known if *BRCA-1*/*BRCA-2* test was performed, ***or*** Not applicable (***cervix***)

INITIAL COURSE OF TREATMENT

Treatment of cancer usually involves a combination of surgery, radiotherapy, chemotherapy, hormones or other systemic therapies, often given over several years.

Population-based registries cannot record the full details of all treatments for all patients, although some registries link their data to more detailed clinical audit data. Many registries record whether a particular type of treatment was given, and the date it was given. Such data should enable international comparison of the distribution of treatment modalities for each cancer, and the interval between diagnosis and first treatment, by age and calendar period.

Stage-specific survival in each jurisdiction may be examined alongside the proportion of patients who received each type of treatment, to help interpretation of international survival comparisons.

We would like to capture some details of the first course of each treatment, e.g. the surgical procedure(s) performed, radiotherapy fractionation, or chemotherapy cycles. The first course of treatment is usually performed within six months of diagnosis.

Treatment for recurrence or progression should not be coded.

Variable 61 Emergency presentation — Optional variable

Numeric variable, one digit.

The risk of emergency presentation varies by age and deprivation group.[33] Diagnosis following an emergency presentation is associated with lower survival. [34]

Code	Meaning
1	The patient ***presented*** as an emergency
2	The patient ***did not present*** as an emergency
9	Not known if the patient presented as an emergency

Variable 62 Cancer-directed surgery — Core variable

Numeric variable, one digit.

This variable is intended to capture evidence on whether ***any surgical procedure*** was performed to treat or manage the patient's cancer. Cancer-directed surgery includes any procedure **done with therapeutic intent**, whether it is major *or* minor, such as breast-conserving surgery, mastectomy or hysterectomy. Removal of a suspicious lesion during a diagnostic examination may also have therapeutic intent, if pathology later confirms the lesion was malignant.

The ***first therapeutic surgical procedure*** may be of curative or palliative intent, depending on the stage of disease at diagnosis.

This variable should ***not include*** any procedure performed **only for diagnostic purposes**, such as needle biopsy, fine needle aspiration, bone marrow biopsy or cytological smear.

Code	Meaning
1	Cancer-directed surgery ***was performed***
2	Cancer-directed surgery ***was not performed***
9	Not known if cancer-directed surgery was performed

Variables 63-65 Date of first cancer-directed surgery **Core variable**

This is the date on which the ***first*** cancer-directed surgical procedure was performed.

Even if re-operation was required within 3 months of the first cancer-directed surgery, please code the date of the ***first*** cancer-directed surgery, not the date of the definitive/last surgery.

If you decide not to provide this date, please code it as missing (99,99,9999) in ***all tumour records***.

Variable 63 Day of first cancer-directed surgery **Core variable**

Numeric variable, one or two digits.

1-31 = The day of the date of first cancer-directed surgery
99 = The day of this date is not known

Variable 64 Month of first cancer-directed surgery **Core variable**

Numeric variable, one or two digits.

1-12 = The month of the date of first cancer-directed surgery
99 = The month of this date is not known

Variable 65 Year of first cancer-directed surgery **Core variable**

Numeric variable, four digits.

YYYY = The year of the date of first cancer-directed surgery
9999 = The year of this date is not known

Variable 66 Type of first cancer-directed surgery **Core variable**

Numeric variable, one digit.

For breast cancer, re-operation (re-excision or mastectomy) may be performed within 3 months of breast-conserving surgery (BCS) if the pathological margin is small, or positive for tumour.[35] The proportion of re-operations varies by surgeon and hospital volume: population-based studies suggest a range of 17–35%.

If mastectomy was required after BCS, code the mastectomy.

Code	Meaning
1 =	Breast-conserving surgery (***breast***) *Use this code for lumpectomy, quadrantectomy, partial mastectomy or segmental mastectomy*
2 =	Mastectomy (***breast***) *Use this code for simple, skin-sparing or nipple-sparing mastectomy* *Use this code also for mastectomy following BCS*
3 =	Simple hysterectomy (***cervix***)
4 =	Radical hysterectomy (***cervix***)
5 =	Bilateral salpingo-oophorectomy and hysterectomy (***ovary***)
6 =	Cytoreductive surgery (***ovary***) *Must include at least lymphadenectomy and omentectomy*
7 =	Cancer-directed surgery ***was not*** performed
9 =	Type of surgery performed is not known

Variable 67 Place of first cancer-directed surgery **Optional variable**

Numeric variable, one digit.

For all **index tumours**, treatment centralised in higher volume/specialised centres is associated with enhanced adherence to guidelines, reduced variation in treatment and better outcomes,[36-38] although these results are still controversial.[39]

Code	Meaning
1 =	General hospital
2 =	Teaching/University hospital
3 =	Specialised cancer centre or cancer hospital
9 =	Place of first cancer-directed surgery is not known, ***or*** Not applicable (cancer-directed surgery ***not*** performed)

Variable 68 **Lymphadenectomy** **Optional variable**

Numeric variable, one digit.

Lymphadenectomy is a surgical procedure in which most or all of the lymph nodes that drain the tumour bed are removed, and checked for malignant invasion. It may be performed for breast cancer (axillary lymphadenectomy)[13,14] and cervical cancer.[15,16]

For ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Lymphadenectomy ***was*** performed
2	=	Lymphadenectomy ***was not*** performed
9	=	Not known if lymphadenectomy was performed, ***or*** Not applicable (***ovary***)

Variables 69-71 **Date of lymphadenectomy** **Optional variable**

This is the date on which lymphadenectomy was performed.

If lymphadenectomy was not performed, or if you decide not to provide this date, please code it as missing (99,99,9999) in ***all tumour records***.

Variable 69 **Day of lymphadenectomy** **Optional variable**

Numeric variable, one or two digits.

1-31	=	The day of the date of lymphadenectomy
99	=	The day of this date is not known

Variable 70 **Month of lymphadenectomy** **Optional variable**

Numeric variable, one or two digits.

1-12	=	The month of the date of lymphadenectomy
99	=	The month of this date is not known

Variable 71 **Year of lymphadenectomy** **Optional variable**

Numeric variable, four digits.

YYYY	=	The year of the date of lymphadenectomy
9999	=	The year of this date is not known

Variable 72 Radiotherapy **Core variable**

Numeric variable, one digit.

This variable is intended to capture evidence on whether ***any form of ionising radiation was used to treat*** or manage the patient's cancer. This includes external beam radiation from any source and brachytherapy (local implantation of radio-isotopes). Radiotherapy may be pre-operative (neo-adjuvant), intra-operative, or post-operative (adjuvant), depending on the site and stage of disease.

The ***first radiotherapy procedure*** may be of curative or palliative intent, depending on the stage of disease at diagnosis.[13-16]

This variable should ***not include*** any form of radio-diagnostic imaging.

This information will be collected **only for women diagnosed with breast or cervical cancer.** For women diagnosed with ovarian cancer, please use code 9 in ***all tumour records***.

Code		Meaning
1	=	Radiotherapy ***was given***
2	=	Radiotherapy ***was not given***
9	=	Not known if radiotherapy was given, ***or*** Not applicable (***ovary***)

Variables 73-75 Date of first radiotherapy **Core variable**

This is the date on which the ***first*** therapeutic radiotherapy procedure was started. This would normally be within six months of a cancer diagnosis.

If radiotherapy was not performed, or if you decide not to provide this date, or for women diagnosed with ovarian cancer, please code the date as missing (99,99,9999) in ***all tumour records***.

Variable 73 Day of first radiotherapy **Core variable**

Numeric variable, one or two digits.

1-31	=	The day of the date of first radiotherapy
99	=	The day of this date is not known

Variable 74 Month of first radiotherapy **Core variable**

Numeric variable, one or two digits.

1-12	=	The month of the date of first radiotherapy
99	=	The month of this date is not known

Variable 75 Year of first radiotherapy **Core variable**

Numeric variable, four digits.

YYYY	=	The year of the date of first radiotherapy
9999	=	The year of this date is not known

Variable 76 **Total radiotherapy dose (Gy)** **Optional variable**

Numeric variable, one or two digits.

Total radiotherapy dose should be reported in **Gray**, as an integer. The Gray is the international system (SI) unit of radiation energy absorbed per unit mass of tissue (1 Gray = 1 Joule/kilogram). The Gray has replaced the rad: 1 Gy = 100 rad.

In 2011, the American Society for Radiation Oncology (ASTRO) endorsed the use of hypo-fractionation (i.e. larger doses of radiation in fewer treatment sessions), especially in the treatment of breast cancer.[40]

For women diagnosed with ovarian cancer, please use code 99 in ***all tumour records***.

Code		Meaning
1-90	=	Total radiotherapy dose in Gray
99	=	Not applicable (radiotherapy was not performed), ***or*** The total radiotherapy dose is not known, ***or*** Not applicable (***ovary***)

Variable 77 **No. of radiotherapy fractions** **Optional variable**

Numeric variable, one or two digits.

Together with the total radiotherapy dose (variable 76), the number of radiotherapy fractions can help understand whether a conventional or hypo-fractionated regime was used.

For women diagnosed with ovarian cancer, please use code 99 in ***all tumour records***

Code		Meaning
1-90	=	Number of radiotherapy fractions
99	=	The number of radiotherapy fractions is not known, ***or*** Not applicable (radiotherapy was not performed), ***or*** Not applicable (***ovary***)

Variable 78 Chemotherapy **Core variable**

Numeric variable, one digit.

Chemotherapy may be given pre-operatively (neo-adjuvant chemotherapy) to reduce tumour size, or post-operatively (adjuvant chemotherapy) to treat residual disease or prevent recurrence, depending on the site and stage of the disease.

The ***first chemotherapy*** may be of curative or palliative intent, depending on the stage of disease at diagnosis.[13-16]

Chemotherapy may be given orally, or via an intravenous, intrathecal or intraperitoneal route.

Code		Meaning
1	=	Chemotherapy ***was given***
2	=	Chemotherapy ***was not given***
9	=	Not known if chemotherapy was given

Variables 79-81 Date of first chemotherapy **Core variable**

This is the date on which the ***first*** chemotherapy was started.

If chemotherapy was not given, or if you decide not to provide this date, please code it as missing (99,99,9999) in ***all tumour records***.

Variable 79 Day of first chemotherapy **Core variable**

Numeric variable, one or two digits.

1-31	=	The day of the date of first chemotherapy
99	=	The day of this date is not known

Variable 80 Month of first chemotherapy **Core variable**

Numeric variable, one or two digits.

1-12	=	The month of the date of first chemotherapy
99	=	The month of this date is not known

Variable 81 Year of first chemotherapy **Core variable**

Numeric variable, four digits.

YYYY	=	The year of the date of first chemotherapy
9999	=	The year of this date is not known

Variable 82 No. of chemotherapy cycles **Optional variable**

Numeric variable, one digit.

Chemotherapy is given in successive cycles,[13-16] e.g. every day for one week followed by three weeks with no chemotherapy. These four weeks constitute one cycle. The length of a cycle typically ranges between two and six weeks.

Code		Meaning
1-90	=	Number of chemotherapy cycles
99	=	Number of chemotherapy cycles is not known, ***or*** Not applicable (chemotherapy was not performed)

Variable 83 **Endocrine treatment** **Core variable**

Numeric variable, one digit.

Endocrine treatment (e.g., aromatase inhibitors, tamoxifen) is the main therapy for hormone receptor-positive (ER+) breast cancer. It may be given orally, or via an intramuscular route.

Endocrine treatment may be given pre-operatively (neo-adjuvant) to reduce tumour size, or post-operatively (adjuvant) to treat residual disease or prevent recurrence, depending on the site and stage of the disease.

The ***first endocrine treatment*** may be of curative or palliative intent, depending on the stage of disease at diagnosis.[13,14]

For breast cancer only. For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	Endocrine treatment ***was given***
2	=	Endocrine treatment ***was not given***
9	=	Not known if endocrine treatment was given, **or** Not applicable (***cervix, ovary***)

Variable 84 **Year of first endocrine treatment** **Core variable**

Numeric variable, four digits.

For breast cancer only. This is the year in which the ***first*** endocrine treatment was started.

If endocrine treatment was not given, or if you decide not to provide this date, please code it as missing (9999) in ***all tumour records***.

YYYY	=	The year of the date of first endocrine treatment
9999	=	The year of this date is not known, ***or*** Not applicable (***cervix, ovary***)

Variable 85 **Anti-HER-2 monoclonal antibody treatment** **Core variable**

Numeric variable, one digit.

Anti-HER-2 monoclonal antibody treatment (e.g. trastuzumab, lapatinib, pertuzumab) is a key therapy for HER-2-positive breast cancer.[13,14]

It may be given pre-operatively (neo-adjuvant) to reduce tumour size, or post-operatively (adjuvant) to treat residual disease or prevent recurrence, depending on the site and stage of the disease.

The ***first anti-HER-2 treatment*** may be of curative or palliative intent, depending on the stage of disease at diagnosis.

For women diagnosed with cervical or ovarian cancer, please use code 9 in ***all tumour records.***

Code		Meaning
1	=	Anti-HER-2 treatment ***was given***
2	=	Anti-HER-2 treatment ***was not given***
9	=	Not known if anti-HER-2 was given, ***or*** Not applicable (***cervix, ovary***)

Variables 86-88 **Date of first anti-HER-2 treatment** **Core variable**

This is the date on which the ***first*** anti-HER-2 treatment was started.

If anti-HER-2 treatment was not performed, or for women diagnosed with cervical or ovarian cancer, please code the date as missing (99,99,9999) ***in all tumour records***.

Variable 86 **Day of first anti-HER-2 treatment** **Core variable**

Numeric variable, one or two digits.

1-31	=	The day of date of first anti-HER-2 treatment
99	=	The day of this date is not known

Variable 87 **Month of first anti-HER-2 treatment** **Core variable**

Numeric variable, one or two digits.

1-12	=	The month of date of first anti-HER-2 treatment
99	=	The month of this date is not known

Variable 88 **Year of first anti-HER-2 treatment** **Core variable**

Numeric variable, four digits.

YYYY	=	The year of date of first anti-HER-2 treatment
9999	=	The year of this date is not known

Variable 89 **Clinical trial** **Optional variable**

Numeric variable, one digit.

This will enable us to assess whether women recruited to a clinical trial have higher odds of being treated according to international guidelines, and whether they have higher survival, than those who are not.

Code		Meaning
1	=	The woman ***was recruited*** to a clinical trial
2	=	The woman ***was not recruited*** to a clinical trial
9	=	Not known if the woman was recruited to a clinical trial

SOCIO-ECONOMIC STATUS

Access to care may be associated with socio-economic status.[41] The measures of socio-economic status differ between countries.[42] We propose to collect the following variables, where available (variables 90-99):

Variable 90 Education — Optional variable

Numeric variable, one digit.

Code		Meaning
1	=	No education
2	=	Primary education
3	=	Secondary education
4	=	Tertiary education (Bachelor, Master, PhD)
9	=	Educational status not known

Variable 91 No. of persons living in the same household — Optional variable

Numeric variable, one digit.

Code		Meaning
1	=	One person (the patient lives alone)
2	=	Two persons
3	=	Three persons
4	=	Four persons
5	=	Five persons
6	=	Six persons or more
9	=	Number of persons in household not known

Variable 92 Home ownership — Optional variable

Numeric variable, one digit.

Home ownership is recognised in many countries as a measure of wealth or socio-economic position.

Code		Meaning
1	=	The patient ***owned*** her own home (alone or jointly)
2	=	The patient ***did not own*** her own home
9	=	Not known if the patient owned her own home

Variable 93 Water supply — Optional variable

Numeric variable, one digit.

Code		Meaning
1	=	Water accessible ***in own home***
2	=	Water ***not*** accessible in own home
9	=	Not known if water accessible in own home

Variable 94 **Electricity supply** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Electricity accessible ***in own home***
2	=	Electricity ***not*** accessible in own home
9	=	Not known if electricity accessible in own home

Variable 95 **Sanitation** **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Sanitation facility available ***in own home***
2	=	Sanitation facility ***shared*** with other household(s)
3	=	No sanitation facility
9	=	Sanitation access not known

Variable 96 **Socio-economic group** **Optional variable**

Numeric variable, one digit.

Detailed information on socio-economic status at individual level is rarely available, so area- and census tract-based measures of socio-economic group are usually used to investigate health disparities.[43] These measures are usually derived from a routinely available socio-economic index for the patient's area of residence at diagnosis. The distribution of the index can then be divided into quintiles.

Code		Meaning
1	=	Quintile 1 (Most affluent)
2	=	Quintile 2
3	=	Quintile 3
4	=	Quintile 4
5	=	Quintile 5 (Most deprived)
9	=	Socio-economic group not available

Variable 97 **Marital status** **Optional variable**

Numeric variable, one digit.

We will use the term "married" to include women living with a partner, even if not married.

Code		Meaning
1	=	Married (or partnered)
2	=	Not married (or partnered)
9	=	Marital status not known

Variable 98 Employment status **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	Employed
2	=	Unemployed
3	=	Retired
9	=	Employment status not known

Variable 99 Health insurance status **Optional variable**

Numeric variable, one digit.

In many countries, health insurance status may determine whether a patient has access to health care.

If you wish to use alternative categories, please contact us **before submitting your data.**

Code		Meaning
1	=	Private
2	=	Public
3	=	Private and public
4	=	Uninsured
9	=	Health insurance status not known

FAMILY HISTORY

The association between a family history of breast or ovarian cancer and the outcome for the patient is still under debate.[44,45]

We are aware that information on family history may be under-reported. However, if relevant data are available, we suggest the following (variables 100-101):

Variable 100 Breast or ovarian cancer in female first-degree relatives **Optional variable**

Numeric variable, one digit.

In this context, first-degree relatives include the patient's mother and any sisters or daughters.

Code		Meaning
1	=	***Absence*** of breast or ovarian cancers in female first-degree relatives
2	=	***Presence*** of breast or ovarian cancers in female first-degree relatives
9	=	Family history not known

Variable 101 No. of affected relatives **Optional variable**

Numeric variable, one digit.

Code		Meaning
1	=	One female first-degree relative with breast or ovarian cancer
2	=	Two female first-degree relatives with breast or ovarian cancer
3	=	Three or more female first-degree relatives with breast or ovarian cancer
4	=	***No*** female first-degree relative with breast or ovarian cancer
9	=	Number of female first-degree relatives with breast or ovarian cancer not known

LIFESTYLE

Variable 102 Alcohol consumption — Optional variable

Numeric variable, one digit.

Alcohol is a well-established risk factor for the development of some cancers, including breast cancer. Little is known about how alcohol use affects cancer treatment delivery.[46] The association between alcohol consumption and cancer survival is also debated.[47,48]

Code		Meaning
1	=	Never drinker
2	=	Former drinker (***see note below***)
3	=	Current drinker
9	=	Drinking status is not known

Note: please use the definition of "former drinker" that is used in your territory.

Variable 103 Smoking habit — Optional variable

Numeric variable, one digit.

Smoking after a cancer diagnosis is associated with several risks, including worse tolerance of treatment, a higher risk of treatment failure or a second primary tumour, and a poorer quality of life.[49]

Code		Meaning
1	=	Never smoker
2	=	Former smoker (***see note below***)
3	=	Current smoker
9	=	Smoking habit is not known

Note: please use the definition of "former smoker" that is used in your territory.

Variable 104 Height (cm) — Optional variable

A patient's height is required to determine the surface area to calculate the dose of some drugs, as well as to determine the Body Mass Index.

Numeric variable, two or three digits – integer.

90-220	=	Height at diagnosis (in centimetres)
999	=	Height at diagnosis is not known

Variable 105 Weight (kg) — Optional variable

Numeric variable, two or three digits – integer, to the nearest whole kilogram.

A patient's weight is required to calculate the dose of some drugs, as well as to determine the Body Mass Index.

35-200	=	Weight at diagnosis (in kilograms)
999	=	Weight at diagnosis is not known

CONCURRENT MEDICAL CONDITIONS

The co-existence of cancer and other medical conditions (co-morbidities) has implications for treatment decisions and survival. For example, diabetic patients may have cardiovascular, renal or neuropathic complications that may influence the choice of chemotherapy.[50]

Many methods for assessing comorbidity are available; no gold standard approach exists.[51]

The Charlson Comorbidity Index is widely used, but it was originally developed for patients with cardiovascular disease. We prefer to collect data on a few specific major disease groups that may affect the choice of cancer treatment.

Comorbid conditions are generally defined as those first recorded in a time window from six years to six months preceding the cancer diagnosis.[52] **Please contact us** if a different definition is used in your territory.

Variable 106 Cardiovascular conditions — Optional variable

Alphanumeric variable, three or four characters.

Please code this variable according to ICD-10.[4]

Please provide the full 4-character ICD-10 code ***without the decimal point (".")*.** For example, acute myocardial infarction, unspecified, should be coded as I219.

Cardiovascular conditions should be coded as follows:

Myocardial infarction	I210-I214, I219; I220-I221, I228-I229; I230-I236, I238, I252
Congestive heart failure	I099, I110, I130, I132, I255, I420–I429, I430-I432, I438, I500-I501, I509, I517, P290
Peripheral vascular disease	I600-I609, I610-I616, I618-I619, I700-I702, I708-I709, I710-I716, I718-I719, I720-I726, I728-I729, I730-I731, I738-I739, I740-I745, I748-I749, I770-I776, I778-I779, I790, I792, K551, K558, K559, R02, Z951, Z955, Z958, Z959
Cerebrovascular disease	G450-G454, G458-G459, G460-G468, I620-I621, I629, I630-I636, I638-I639, I64, I650-I653, I658-I659, I660-I664, I668-I669, I670-I679, I680-I682, I688, I690-I694, I698, H340

Variable 107 Diabetes — Optional variable

Alphanumeric variable, four characters.

Please code this variable according to ICD-10.[4]

Please provide the full 4-character ICD-10 code ***without the decimal point (".")*.** For example, type 1 diabetes mellitus without complications should be coded as E109.

Type of diabetes should be coded as follows:

Diabetes	E100-E109, E110-E119, E120-E129, E130-E139, E140-E149

Variable 108 Previous cancer **Optional variable**

Alphanumeric variable, four characters.

Please code the full 4-character ICD-O-3[2] topography code ***without the decimal point (".")***. Topography codes are the same as in the first revision of ICD-O-3.[6] For example, breast cancer of the upper inner quadrant should be coded as C502.

If more than one previous cancer has been recorded, please code the site of the most recent.

Please use code 9999 if no previous cancer has been recorded.

Variable 109 Renal impairment **Optional variable**

Alphanumeric variable, three or four characters.

Please code this variable according to ICD-10.[4]

Please provide the full 3- or 4-character ICD-10 code ***without the decimal point (".")*.** For example, hypertensive renal disease with renal failure should be coded as I120.

Renal impairment should be coded as follows:

Renal impairment	I120-I129; I130-I131, I139; N000-N009, N010-N019, N020-N029, N030-N039, N040-N049, N050-N059, N070-N079, N110-N111, N118-N119; N140-N144, N170-N172, N178-N179, N180-N185, N189, N19, N250, Q610-Q615, Q618-Q619, Z490-Z492, Z940, Z992

Variable 110 Liver impairment **Optional variable**

Alphanumeric variable, four characters.

Please code this variable according to ICD-10.[4]

Please provide the full 4-character ICD-10 code ***without the decimal point (".")*.** For example, chronic viral hepatitis, unspecified, should be coded as B189.

Mild impairment	B180-B189, K700-K709, K710-K719, K730-K739, K740-K746, K760, K761-K769, Z944
Moderate or severe impairment	B150, B160, B162, B190, I850, I859, I864, I982, K704, K711, K721, K729, K765, K766, K767

Variable 111 History of HIV **Optional variable**

Numerical variable, one digit.

Code	Meaning
1 =	***Absence*** of HIV
2 =	***Presence*** of HIV
3 =	HIV not determined
9 =	Not stated, ***or*** Not available

4 Valid values for each variable

Name of variable	Short description	No. of digits or characters	Valid values (or *range* of valid values)	Value to be used when valid data are missing
	Demographics			
VAR1	Country	4	Use the value provided *	Not allowed
VAR2	Registry	1, 2, or 3	Use the value provided *	Not allowed
VAR3	Person code	Up to 15	Numeric or alphanumeric	Not allowed
VAR4	Tumour code	Up to 10	Numeric or alphanumeric	Not allowed
VAR5	Sex	1	1,2	9
VAR6	Region *(notes 1, 2 below)*	Up to 5	Numeric (see p6)	9 *or* 99999
VAR7	Race/ethnicity *(notes 1, 2 below)*	1 or 2	Numeric (see p7)	99
VAR8	Day of birth	1 or 2	1-31	99
VAR9	Month of birth	1 or 2	1-12	99
VAR10	Year of birth	4	1915 - 2003	9999
VAR11	Day of diagnosis	1 or 2	1-31	99
VAR12	Month of diagnosis	1 or 2	1-12	99
VAR13	Year of diagnosis	4	2015 or a later year [§]	Not allowed
	Follow-up for vital status			
VAR14	Last known vital status	1	1,2,3	9
VAR15	Day of last known vital status	1 or 2	1-31	99
VAR16	Month of last known vital status	1 or 2	1-12	99
VAR17	Year of last known vital status	4	2015 - 2019	9999
	Tumour details			
VAR18	Basis of diagnosis	1	1,2,3,4,5,6,7,8	9
VAR19	Topography	4	Alphanumeric (see p13)	Not allowed
VAR20	Morphology	4	8000-9992	9999
VAR21	Behaviour	1	0,1,2,3,6,9	Not allowed
VAR22	Grade	1	1,2,3,4,8	9
VAR23	CIN grade *(note 3 below)*	1	1,2,3,8	9
VAR24	Multifocality	1	1,2,3	9
VAR25	Laterality *(note 4 below)*	1	1,2,3	9
VAR26	Screen-detected *(note 5 below)*	1	1,2	9
	Stage of disease at diagnosis			
VAR27	Pathological T	1	0,1,2,3,4,8	9
VAR28	Pathological N	1	0,1,2,3	9
VAR29	Pathological M	1	1	9
VAR30	Clinical T	1	0,1,2,3,4,8	9
VAR31	Clinical N	1	0,1,2,3	9
VAR32	Clinical M	1	0,1	9
VAR33	Site of metastasis *(note 6 below)*	1	1,2,3	9
VAR34	SEER Summary Stage 2000	1	0,1,2,3,4,5,7	9
VAR35	Condensed T	1	1,2	9
VAR36	Condensed N	1	1,2	9
VAR37	Condensed M	1	1,2	9
VAR38	FIGO stage *(note 7 below)*	1	0,1,2,3,4,5	9
VAR39	Tumour size (mm)	1, 2, or 3	1-300	999
VAR40	Sentinel lymph node biopsy *(note 5 below)*	1	1,2	9
VAR41	No. of lymph nodes examined	1 or 2	0-98	99
VAR42	No. of lymph nodes involved	1 or 2	0-98	99

Name of variable	Short description	No. of digits or characters	Valid values (or *range* of valid values)	Value to be used when valid data are missing
	Diagnostic procedures			
VAR43	Mammography *(note 6 below)*	1	1,2,3,4	9
VAR44	Tissue diagnosis *(note 5 below)*	1	1,2,3	9
VAR45	Papanicolaou (Pap) test *(note 3 below)*	1	1,2	9
VAR46	HPV test *(note 3 below)*	1	1,2	9
VAR47	Colposcopy *(note 3 below)*	1	1,2	9
VAR48	Chest X-ray	1	1,2	9
VAR49	Abdominal ultrasound	1	1,2	9
VAR50	Scintigraphy	1	1,2	9
VAR51	Skeletal X-ray	1	1,2	9
VAR52	Computerised Tomography (CT) scan	1	1,2	9
VAR53	Trans-vaginal ultrasound *(note 7 below)*	1	1,2	9
	Biomarkers			
VAR54	[O]estrogen receptors *(note 6 below)*	1	1,2,3	9
VAR55	Progesterone receptors *(note 6 below)*	1	1,2,3	9
VAR56	HER-2 based on IHC assay *(note 6 below)*	1	1,2,3,4	9
VAR57	HER-2 based on FISH assay *(note 6 below)*	1	1,2,3,4	9
VAR58	Ki-67 proliferation index	1, 2, or 3	0-100,888	999
VAR59	CA-125 *(note 8 below)*	1	1,2	9
VAR60	*BRCA-1* or *BRCA-2* *(note 4 below)*	1	1,2,3	9
	Initial course of treatment			
VAR61	Emergency presentation	1	1,2	9
VAR62	Cancer-directed surgery	1	1,2	9
VAR63	Day of first cancer-directed surgery	1 or 2	1-31	99
VAR64	Month of first cancer-directed surgery	1 or 2	1-12	99
VAR65	Year of first cancer-directed surgery	4	2015 or a later year [§]	9999
VAR66	Type of first cancer-directed surgery	1	1,2,3,4,5,6,7	9
VAR67	Place of first cancer-directed surgery	1	1,2,3	9
VAR68	Lymphadenectomy *(note 5 below)*	1	1,2	9
VAR69	Day of lymphadenectomy *(note 5 below)*	1 or 2	1-31	99
VAR70	Month of lymphadenectomy *(note 5 below)*	1 or 2	1-12	99
VAR71	Year of lymphadenectomy *(note 5 below)*	4	2015 or a later year [§]	9999
VAR72	Radiotherapy *(note 5 below)*	1	1,2	9
VAR73	Day of first radiotherapy *(note 5 below)*	1 or 2	1-31	99
VAR74	Month of first radiotherapy *(note 5 below)*	1 or 2	1-12	99
VAR75	Year of first radiotherapy *(note 5 below)*	4	2015 or a later year [§]	9999
VAR76	Total radiotherapy dose (Gy) *(note 5 below)*	1 or 2	1-90	99
VAR77	No. of radiotherapy fractions *(note 5 below)*	1 or 2	1-90	99
VAR78	Chemotherapy	1	1,2	9
VAR79	Day of first chemotherapy	1 or 2	1-31	99
VAR80	Month of first chemotherapy	1 or 2	1-12	99
VAR81	Year of first chemotherapy	4	2015 or a later year [§]	9999
VAR82	No. of chemotherapy cycles	1 or 2	1-90	99
VAR83	Endocrine treatment *(note 6 below)*	1	1,2	9
VAR84	Year of first endocrine treatment *(note 6)*	4	2015 or a later year [§]	9999
VAR85	Anti-HER-2 treatment *(note 6 below)*	1	1,2	9
VAR86	Day of first anti-HER-2 treatment *(note 6)*	1 or 2	1-31	99
VAR87	Month of first anti-HER-2 treatment *(note 6)*	1 or 2	1-12	99
VAR88	Year of first anti-HER-2 treatment *(note 6)*	4	2015 or a later year [§]	9999
VAR89	Clinical trial	1	1,2	9

Name of variable	Short description	No. of digits or characters	Valid values (or *range* of valid values)	Value to be used when valid data are missing
	Socioeconomic status			
VAR90	Education	1	1,2,3,4	9
VAR91	No. of persons living in the same household	1	1,2,3,4,5,6	9
VAR92	Home ownership	1	1,2	9
VAR93	Water supply	1	1,2	9
VAR94	Electricity supply	1	1,2	9
VAR95	Sanitation	1	1,2,3	9
VAR96	Socio-economic group	1	1,2,3,4,5	9
VAR97	Marital status	1	1,2	9
VAR98	Employment status	1	1,2,3	9
VAR99	Health insurance status	1	1,2,3,4	9
	Family history			
VAR100	Breast/ovarian cancer in 1st-degree relatives	1	1,2	9
VAR101	No. of affected 1st-degree relatives	1	1,2,3,4	9
	Lifestyle			
VAR102	Alcohol consumption	1	1,2,3	9
VAR103	Smoking habit	1	1,2,3	9
VAR104	Height (cm)	2 or 3	90-220	999
VAR105	Weight (kg)	2 or 3	35-200	999
	Concurrent medical conditions			
VAR106	Cardiovascular conditions	3 or 4	Alphanumeric (see p44)	9999
VAR107	Diabetes	3 or 4	Alphanumeric (see p44)	9999
VAR108	Previous cancer	4	Alphanumeric (see p45)	9999
VAR109	Renal impairment	3 or 4	Alphanumeric (see p45)	9999
VAR110	Liver impairment	3 or 4	Alphanumeric (see p45)	9999
VAR111	History of HIV	1	1,2,3	9

[*] We will send you the code to use for this variable; it should be used ***in all tumour records***

[§] Either 2015 or the later calendar year of incidence for which you will submit data

[1] If you ***do not wish*** to provide data for these variables, please use the "missing value" code (last column) ***in all tumour records***

[2] If you ***do wish*** to provide data for these variables, please contact us first, to agree the codes that you will use

[3] Valid values only for **cervical cancer**; for breast and ovarian cancers, please use the "missing value" code ***in all tumour records***

[4] Valid values only for **breast and ovarian cancer**; for cervical cancer, please use the "missing value" code ***in all tumour records***

[5] Valid values only for **breast and cervical cancer**; for ovarian cancer, please use the "missing value" code ***in all tumour records***

[6] Valid values only for **breast cancer**; for cervical and ovarian cancer, please use the "missing value" code ***in all tumour records***

[7] Valid values only for **cervical and ovarian** cancer; for breast cancer, please use the "missing value" code ***in all tumour records***

[8] Valid values only for **ovarian cancer**; for breast and cervical cancer, please use the "missing value" code ***in all tumour records***

5 REFERENCES

1. Allemani C, Matsuda T, Di Carlo V*, et al.* Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 2018; **391**: 1023-75. http://dx.doi.org/10.1016/S0140-6736(17)33326-3
2. Fritz AG, Percy C, Jack A*, et al.*, eds. International Classification of Diseases for Oncology (ICD-O). 3rd edn. Geneva: World Health Organization; 2000.
3. World Health Organisation. International Classification of Diseases, 1975. Ninth revision. Geneva: WHO; 1977.
4. World Health Organisation. International statistical classification of diseases and related health problems. Tenth revision. Geneva: WHO; 1994.
5. Percy C, Van Holten V, Muir CS, eds. International Classification of Diseases for Oncology (ICD-O). 2nd edn. Geneva: World Health Organisation; 1990.
6. Fritz AG, Percy C, Jack A*, et al.*, eds. International Classification of Diseases for Oncology (ICD-O). First revision of 3rd edn. Geneva: World Health Organisation; 2013.
7. Woods LM, Rachet B, Ellis L*, et al.* Full dates (day, month, year) should be used in population-based cancer survival studies. *Int J Cancer* 2012; **131**: E1120-E4. 10.1002/ljc.27545
8. Wodarz D, Iwasa Y, Komarova NL. On the emergence of multifocal cancers. *J Carcinog* 2004; **3**: 1-13. doi:10.1186/1477-3163-3-13
9. Lu KH, Bell DA, Welch WR*, et al.* Evidence for the multifocal origin of bilateral and advanced human serous borderline ovarian tumors. *Cancer Res* 1998; **58**: 2328-30.
10. Schorge JO, Muto MG, Lee SJ*, et al.* BRCA1-related papillary serous carcinoma of the peritoneum has a unique molecular pathogenesis. *Cancer Res* 2000; **60**: 1361-4.
11. McIlwaine P, Nagar H, McCluggage WG. Multifocal FIGO Stage 1A1 cervical squamous carcinomas have an extremely good prognosis equivalent to unifocal lesions. *Int J Gynecol Pathol* 2014; **33**: 213-7. 10.1097/PGP.0b013e31829040ce
12. Dickie L, Johnson CH, Adams S*, et al.* Solid Tumor Rules. Rockville, MD: National Cancer Institute, 2019. https://seer.cancer.gov/tools/solidtumor/ (accessed 1 Dec 2019).
13. European Society for Medical Oncology. Breast Cancer 2018. Lugano, Switzerland: ESMO; 2018.
14. American Society of Clinical Oncology. Guidelines on Breast Cancer. Alexandria VA: ASCO, 2019. https://www.asco.org/research-guidelines/quality-guidelines/guidelines/breast-cancer (accessed 1 Dec 2019).
15. European Society for Medical Oncology. Gynaecological Malignancies 2018. Lugano, Switzerland: ESMO; 2018.
16. American Society of Clinical Oncology. Guidelines on Gynecological Cancers. Alexandria VA: ASCO, 2019. https://www.asco.org/research-guidelines/quality-guidelines/guidelines/gynecologic-cancer (accessed 1 Dec 2019).
17. Allemani C, Sant M, Weir HK*, et al.* Breast cancer survival in the US and Europe: a CONCORD high-resolution study. *Int J Cancer* 2013; **132**: 1170-81.
18. Sant M, Allemani C, Capocaccia R*, et al.* Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer* 2003; **106**: 416-22.
19. Sankaranarayanan R, Swaminathan R, eds. Cancer survival in Africa, Asia, the Caribbean and Central America (IARC Scientific Publications No 162). Lyon: IARC; 2011.
20. Maringe C, Walters S, Butler J*, et al.* Stage at diagnosis and ovarian cancer survival: evidence from the International Cancer Benchmarking Partnership. *Gynecol Oncol* 2012; **127**: 75-82.
21. Sobin LH, Gospodarowicz M, Wittekind C, eds. TNM Classification of Malignant Tumours. 7th edn. New York: John Wiley & Sons; 2009.
22. Edge SB, Byrd DR, Carducci MA*, et al.*, eds. American Joint Committee on Cancer - cancer staging manual. 7th edn. New York, NY: Springer; 2010.
23. Young JL, Roffers SD, Ries LAG*, et al.* SEER Summary Staging Manual - 2000: Codes and Coding Instructions. NIH Pub. No. 01-4969. Bethesda, MD: National Cancer Institute; 2001.
24. Berrino F, Brown C, Moller T*, et al.* ENCR recommendations: condensed TNM for coding the extent of disease. Lyon: European Network of Cancer Registries, 2002. http://www.encr.eu/images/docs/recommendations/extentofdisease.pdf.
25. Sobin LH, Wittekind C, eds. TNM Classification of Malignant Tumours. 6th edn. New York: John Wiley & Sons; 2002.
26. Benedet JL, Bender H, Jones H*, et al.* FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. *Int J Gynaecol Obstet* 2000; **70**: 209-62.
27. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med* 1985; **312**: 1604-8. 10.1056/NEJM198506203122504
28. Tung NM, Garber JE. BRCA1/2 testing: therapeutic implications for breast cancer management. *Br J Cancer* 2018; **119**: 141-52. 10.1038/s41416-018-0127-5

29. Madariaga A, Lheureux S, Oza AM. Tailoring ovarian cancer treatment: implications of BRCA1/2 mutations. *Cancers* 2019; **11**: 32. https://doi.org/10.3390/cancers11030416
30. Hammond MEH, Hayes DF, Dowsett M*, et al.* American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *Arch Pathol Lab Med* 2010; **134**: 907-22. 10.1043/1543-2165-134.6.907
31. Wolff AC, Hammond MEH, Hicks DG*, et al.* Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 2013; **31**: 3997-4013. 10.1200/jco.2013.50.9984
32. Coates AS, Winer EP, Goldhirsch A*, et al.* Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015; **26**: 1533-46. 10.1093/annonc/mdv221
33. Abel GA, Shelton J, Johnson S*, et al.* Cancer-specific variation in emergency presentation by sex, age and deprivation across 27 common and rarer cancers. *Br J Cancer* 2015; **112 (Suppl 1)**: 129-36. 10.1038/bjc.2015.52
34. Elliss-Brookes L, McPhail S, Ives A*, et al.* Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. *Br J Cancer* 2012; **107**: 1220-6.
35. Monaghan A, Chapinal N, Hughes L*, et al.* Impact of SSO-ASTRO margin guidelines on reoperation rates following breast-conserving surgery. *Am J Surg* 2019; **217**: 862-7. https://doi.org/10.1016/j.amjsurg.2019.01.007
36. Richardson LC, Tian L, Voti L*, et al.* The roles of teaching hospitals, insurance status, and race/ethnicity in receipt of adjuvant therapy for regional-stage breast cancer in Florida. *Am J Public Health* 2006; **96**: 160-6. 10.2105/ajph.2004.053579
37. Matsuo K, Shimada M, Yamaguchi S*, et al.* Association of radical hysterectomy surgical volume and survival for early-stage cervical cancer [Epub ahead of print]. *Obstet Gynecol* 2019; **133**: 1086-98. doi: 10.1097/AOG.0000000000003280
38. White KM, Seale H, Harrison R. Enhancing ovarian cancer care: a systematic review of guideline adherence and clinical variation. *BMC Public Health* 2019; **19**: 296-309. https://doi.org/10.1186/s12889-019-6633-4
39. Siesling S, Tjan-Heijnen VCG, de Roos M*, et al.* Impact of hospital volume on breast cancer outcome: a population-based study in the Netherlands. *Breast Cancer Res Treat* 2014; **147**: 177-84. https://doi.org/10.1007/s10549-014-3075-7
40. Smith BD, Bentzen SM, Correa CR*, et al.* Fractionation for whole breast irradiation: an American Society for Radiation Oncology (ASTRO) evidence-based guideline. *Int J Radiat Oncol Biol Phys* 2011; **81**: 59-68. 10.1016/j.ijrobp.2010.04.042
41. Lyratzopoulos G, Abel GA, Brown CH*, et al.* Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. *Ann Oncol* 2013; **24**: 843-50. 10.1093/annonc/mds526
42. Guillaume E, Pornet C, Dejardin O*, et al.* Development of a cross-cultural deprivation index in five European countries. *J Epidemiol Comm Hlth* 2016; **70**: 493-9. 10.1136/jech-2015-205729
43. Yu M, Tatalovich Z, Gibson JT*, et al.* Using a composite index of socioeconomic status to investigate health disparities while protecting the confidentiality of cancer registry data. *Cancer Causes Control* 2014; **25**: 81-92. 10.1007/s10552-013-0310-1
44. Anderson K, Thompson PA, Wertheim BC*, et al.* Family history of breast and ovarian cancer and triple negative subtype in hispanic/latina women. *SpringerPlus* 2014; **3**: 727. https://doi.org/10.1016/j.ijrobp.2010.04.042
45. Ji J, Försti A, Sundquist J*, et al.* Survival in ovarian cancer patients by histology and family history. *Acta Oncol* 2008; **47**: 1133-9. 10.1080/02841860701784544
46. LoConte NK, Brewster AM, Kaur JS*, et al.* Alcohol and cancer: a statement of the American Society of Clinical Oncology. *J Clin Oncol* 2017; **36**: 83-93. 10.1200/JCO.2017.76.1155
47. Allemani C, Berrino F, Krogh V*, et al.* Do pre-diagnostic drinking habits influence breast cancer survival? *Tumori* 2011; **97**: 142-8.
48. Newcomb PA, Kampman E, Trentham-Dietz A*, et al.* Alcohol consumption before and after breast cancer diagnosis: associations with survival from breast cancer, cardiovascular disease, and other causes. *J Clin Oncol* 2013; **31**: 1939-46. 10.1200/JCO.2012.46.5765
49. Jassem J. Tobacco smoking after diagnosis of cancer: clinical aspects. *Trans Lung Cancer Res* 2019; **8 (Suppl 1)**: 50-8. http://dx.doi.org/10.21037/tlcr.2019.04.01
50. Jacob P, Chowdhury TA. Management of diabetes in patients with cancer. *QJM* 2014; **108**: 443-8. 10.1093/qjmed/hcu218
51. Williams GR, Mackenzie A, Magnuson A*, et al.* Comorbidity in older adults with cancer. *J Geriatr Oncol* 2016; **7**: 249-57. https://doi.org/10.1016/j.jgo.2015.12.002
52. Maringe C, Fowler H, Rachet B*, et al.* Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities. *PLoS One* 2017; **12**: e0172814. doi:10.1371/journal.pone.0172814