Standard Procedure Manual

for Population-based Cancer Registries in sub Saharan Africa

Version 5

Nov 2024

Edited by A.M. Finesse, N. Somdyala, E. Chokunonga, D.M. Parkin on behalf of African Cancer Registry Network

Prama House, 267 Banbury Road, Oxford OX2 7HT, United Kingdom

Contents

List of Figures

List of Tables

ACKNOWLEDEGEMENTS

Editors are grateful to the following individuals who made invaluable contributions to this manual:

Dr Ariana Znaor and Mr Morten Ervik (IARC) for their many helpful suggestions and additions.
Dr Eva Kantelhardt (University of Halle) for her advice.
Directors of the institutions in which the editors are based.
Mrs Biying Liu for editing of the final manual.

Editors would also like to thank these organisations for their tremendous support.

African Cancer Registry Network (AFCRN)
International Agency for Research on Cancer (IARC)
American Cancer Society (ACS)
Union for International Cancer Control (UICC)
Vital Strategies

GLOSSARY

AFCRN	African Cancer Registry Network
AJCC	American Joint Committee on Cancer
CI5	Cancer Incidence in Five Continents
EARN	East African Cancer Registry Network
IACR	International Association of Cancer Registries
IARC	International Agency for Research on Cancer
ICD-O	International Classification of Diseases for Oncology
INCTR	International Network for Cancer Treatment and Research
NAACCR	North American Association of Central Cancer Registries
NGO	Non-governmental organization
NOS	Not otherwise specified
PBCR	Population-based cancer registry
UICC	Union for International Cancer Control
WHO	World Health Organization

1. BACKGROUND

Cancer is public health problem in both developed and developing countries as a result of the increase in life expectancy, changes in diet, lifestyle and other factors. To help address this growing burden in some African countries, the Cancer Registry Programme of the International Network for Cancer Treatment and Research (INCTR) established the East African Cancer Registry Network (EARN) which later became the African Cancer Registry Network (AFCRN) in 2012. The overall aim of the network is to improve the effectiveness of cancer registration and surveillance in sub–Saharan Africa by providing expert evaluation of current problems and technical support to remedy identified barriers, with long term goals of strengthening health systems and creating research platforms for the identification of problems, priorities and targets for intervention. The AFCRN serves as the "regional hub" of the Global Initiative for Cancer Registration (GICR) of the International Agency for Research on Cancer (IARC). Sub-Saharan African countries, like countries in other regions, urgently need these data for cancer control planning, intervention programmes and subsequently for the use by Health Ministries, policy makers, researchers, clinicians, non-governmental organization (NGOs) and other stakeholders. All countries should have at least one population-based cancer registry, for the purposes of setting priorities, targets for interventions, and monitoring success of cancer control.

2. INTRODUCTION

This manual provides a model, or template, for a Standard Procedure Manual for cancer registries of sub-Saharan Africa. It provides a guide for registry staff in the processes needed to register cancer cases (case finding, abstracting, coding, data entry and storage). Most of the definitions and coding schemes are from international handbooks and guidelines, particularly those published by the International Agency for Research on Cancer (IARC), the International Association of Cancer Registries (IACR), and the Union for International Cancer Control (UICC).

Since every registry is different, some sections are provided only as instructions, guidelines, or examples. **These are shown in blue font, *with* or without *italics***. Each registry will have to customize these sections for their own use. The items of information that are described in the manual (for abstracting from medical records, coding, and entering on the computer) correspond to the AFCRN minimum data set. It is recognized that some registries will wish to collect more than this (for example, patient occupation, co-morbidity, morphological grade).

At the present time (2024) the majority of African cancer registries still use traditional manual procedures to collect information on cancer cases – that is, entering the required items of information onto a paper form, which is then coded, and entered into a computerized database. In a few instances, the data sources being used (hospitals, laboratories, death registration systems) have their own computer database, which can be used to directly transfer the required data to the registry. However, in this manual, the emphasis is on the manual methods, with reference, where appropriate, to more automated techniques.

3. CASE FINDING

3.1 SOURCES OF INFORMATION (where to collect data)

All possible sources of cancer information for the registry should be identified. The main sources of information are hospitals, pathology laboratory reports, radiology departments, medical records, death certificates, postmortem/autopsy reports, and radiotherapy and oncology units. However, a registry may also cover private clinics and general practitioners, hospices and screening programmes to ensure completeness.

A list should be prepared for each of the data sources in a hospital to facilitate case finding. The registry should maintain a log book indicating sources covered and when they are covered. Ideally a registry should have a data collection time table indicating frequency of visits to various sources. This monitors the completeness of case finding.

Data sources in hospitals

Use all available sources to ensure the collection of most cases, including those diagnosed clinically (no histo-pathological confirmation), and those that were diagnosed histologically.

I) Medical records: out-patients and in-patients records, admission and discharge forms or books.
- *If there are discrepancies between the diagnosis on admission and that on discharge, the discharge diagnosis is preferred.*

II) Pathology records: pathology reports, autopsy reports and cytology and haematology reports

III) Radiology records: CT scan reports, MRI reports, Ultra sound reports and mammography reports

IV) Radiotherapy department

V) Oncology department

VI) Mortuary register

VII) Disease index (in medical records department)
- *This may be a card index or a hospital information system on computer. However, the disease index may be incomplete, so the registry staff should attempt to cover all available data sources in order to achieve complete registration.*

Other sources

These include: hospices (VERY useful), health insurance systems, screening programmes and central registries.

Death certificates

Are death certificates mentioning CAUSE of death available?
- *In the hospitals?*
- *In civil registration (births, marriages, deaths) offices?*

If they exist, they should be used as a source if at all possible.

3.2 REPORTABLE LIST

What cancer cases should be reported to the registry?

All **cancer cases**, in persons who are **resident** in the ***[Registry Area]***[1] diagnosed since ***[date]*** must be reported to the registry.

[1] *The registry may decide to register ONLY cancer cases normally resident in the "target population" of the registry, OR to record all cases traced in the sources of information being used, and sorting residents from non-residents at the time of analysis of the registry database*

Cancer cases include:

- Cases considered as malignant in the morphology section of the International Classification of Diseases for Oncology (ICD-O); behaviour code-3 should be reported to the registry.
- Benign tumours and tumours of uncertain behaviour of the brain (behaviour code-0, 1, 2).
- **Squamous intraepithelial lesions of the uterine cervix**
  Squamous intraepithelial lesions (SILs) of the uterine cervix, also known as cervical intraepithelial neoplasia (CIN), are proliferations of squamous cells driven by HPV infection, and are divided into low-grade SILs (LSILs) and high-grade SILs (HSILs). CANCER REGISTRIES SHOULD RECORD ALL CASES OF HSIL, also known as Cervical intra-epithelial neoplasia grade 2 or 3 (CIN 2, CIN 3); Moderate/severe squamous dysplasia, and Squamous cell carcinoma in situ

Cancers in metastatic sites (for example, lymph nodes) are common, especially in pathology reports. These cases should be registered with the tumour site ("topography") = the site of the primary tumour. If this is not known, register as "unknown primary site" (see section 4.4.3, page 15).

ALL cancer patients, no matter how they were diagnosed, must be reported, including patients with a clinical diagnosis of cancer based only on clinical judgment.

All cancer cases diagnosed at autopsy must be reported.

Patients with active disease and a history of cancer must be reported.

Unclear terms may be found in case notes or laboratory reports, when the physicians are not sure about the behaviour of a tumour (usually when no histological examination has been done). The following table provides some guidance as which should, or should not, be registered.

Accept As Cancer	Not Cancer (Not Reportable)
• Apparently (malignant) • Presumed (malignant) • Compatible with (malignancy) • Probable (malignancy) • Suspect or suspected (malignancy) • Suspicious of (malignancy) • Most likely (malignant) • Consistent with (malignancy)	• "Rule out" • "Equivocal" • "Possible" • "Suggestive" • "Questionable" • "Very close" • " Approaching" • " Encroaching upon"

In cases where the diagnosis remains doubtful, the details should be abstracted but kept in a pending file.

Resident

A person is normally considered a resident in the registry area if they have lived there for *3/4/6/12 months.* Temporary residents – for example, those coming into the area for medical treatment (often lodging with relatives) – must be excluded.

HOWEVER, it is difficult to apply this *x month* rule unless the cancer patient is interviewed.

Normally, we rely on "place of residence" as recorded by staff in the admissions/medical record office. These staff should be encouraged to enquire about the true/usual residence (and not just the temporary "contact" address).

4. ABSTRACTING

Abstracting is the process of extracting from various source documents the information needed to make a registration of cancer.

4.1 REGISTRATION/NOTIFICATION FORM

In paper-based abstracting, the information is entered onto a REGISTRATION/NOTIFICATION FORM (Fig 4.1)

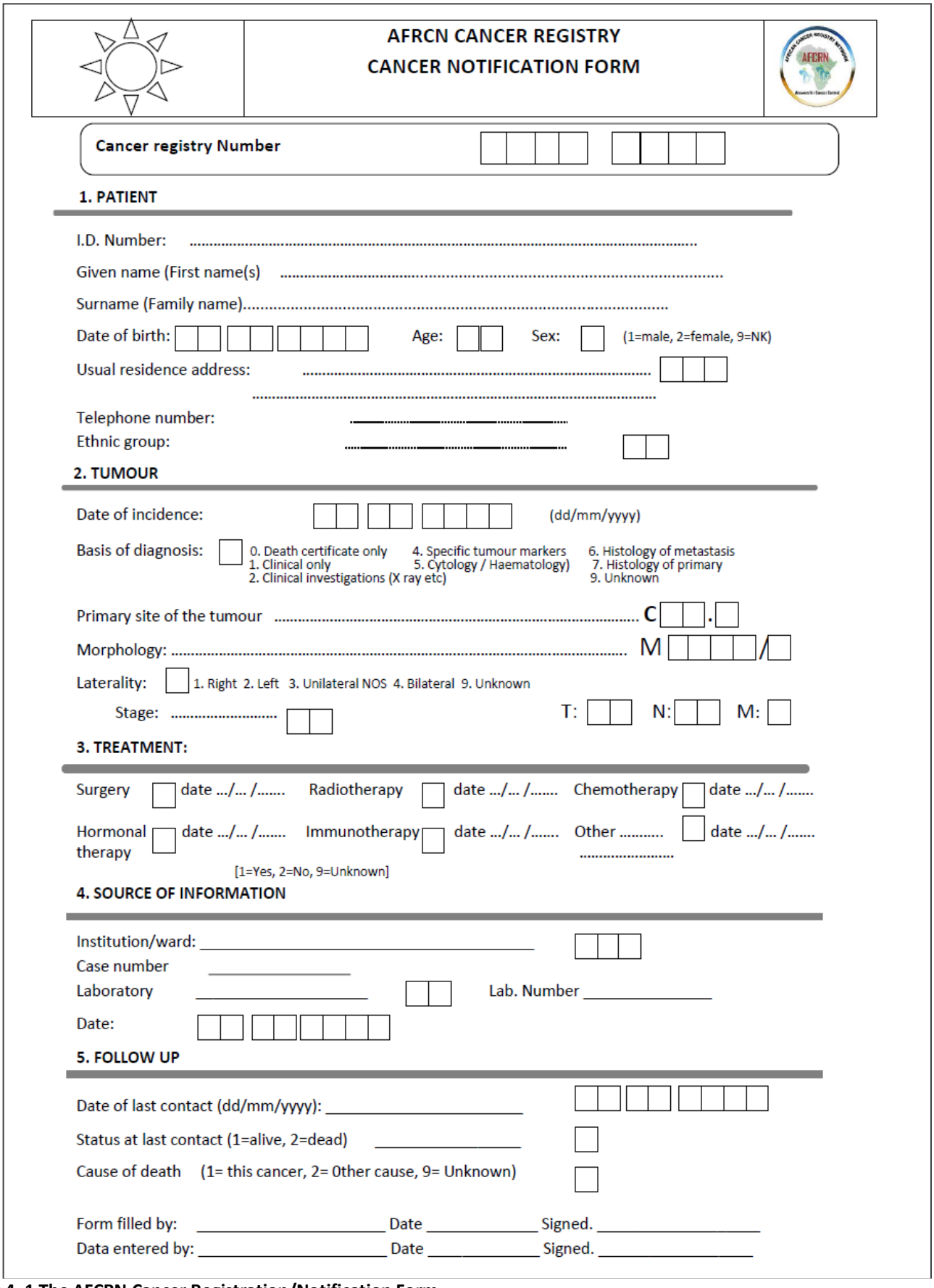
Check before abstraction onto the form

- Is the diagnosis reportable?
- When was the incidence date?
    - Was the incidence date on or after [reference date of the registry]?

COMPLETE A REGISTRATION/ NOTIFICATION FORM EACH TIME A REPORTABLE CANCER CASE IS FOUND IN ANY OF THE INFORMATION SOURCES YOU ARE USING

The different variables that are collected are divided into:

- **Mandatory variables**: These MUST be completed, or a record cannot be confirmed in CanReg. They are: patient names, usual residential address, age, sex, incidence date, most valid basis of diagnosis, primary site, histological type, and behaviour
- **Optional variables**: Telephone number, ethnic group, laterality, stage, TNM, grade, treatment, follow-up status (date of last contact, vital status)



AFRCN CANCER REGISTRY
CANCER NOTIFICATION FORM

Cancer registry Number

1. PATIENT

I.D. Number: .....

Given name (First name(s) .....

Surname (Family name).....

Date of birth: Age: Sex: (1=male, 2=female, 9=NK)

Usual residence address: .....

.....

Telephone number: .....

Ethnic group: .....

2. TUMOUR

Date of incidence: (dd/mm/yyyy)

Basis of diagnosis: 0. Death certificate only 1. Clinical only 2. Clinical investigations (X ray etc) 4. Specific tumour markers 5. Cytology / Haematology) 6. Histology of metastasis 7. Histology of primary 9. Unknown

Primary site of the tumour ..... C..

Morphology: ..... M...../

Laterality: 1. Right 2. Left 3. Unilateral NOS 4. Bilateral 9. Unknown

Stage: ..... T: N: M:

3. TREATMENT:

Surgery date .../... /..... Radiotherapy date .../... /..... Chemotherapy date .../... /.....

Hormonal therapy date .../... /..... Immunotherapy date .../... /..... Other ..... date .../... /.....

[1=Yes, 2=No, 9=Unknown]

4. SOURCE OF INFORMATION

Institution/ward: _____

Case number _____

Laboratory _____ Lab. Number _____

Date:

5. FOLLOW UP

Date of last contact (dd/mm/yyyy): _____

Status at last contact (1=alive, 2=dead) _____

Cause of death (1= this cancer, 2= 0ther cause, 9= Unknown)

Form filled by: _____ Date _____ Signed. _____

Data entered by: _____ Date _____ Signed. _____

Fig 4. 1 The AFCRN Cancer Registration/Notification Form

4.2 DEATH CERTIFICATES

For death certificates, look at the section related to "Cause of Death", page 24. A typical section of a death certificate – the part where causes of death are written – is shown as Fig 4. 2. Note that the doctor can write several medical conditions – those that lead to the death, and those that might have contributed to it.

COMPLETE A CASE REGISTRATION/NOTIFICATION FORM FOR PERSONS WITH CANCER MENTIONED ANYWHERE ON THE CERTIFICATE

Cause of death		Approximate interval between onset and death
I Disease or condition directly leading to death*	(a)	.....
	due to (or as a consequence of)	
Antecedent causes Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last	(b)	.....
	due to (or as a consequence of)	
	(c)	.....
II Other significant conditions contributing to the death, but not related to the disease or condition causing it	.....	.....
	.....	.....
**This does not mean the mode of dying, e.g. heart failure, respiratory failure. It means the disease, injury, or complication that caused death.*		

Fig 4. 2 International Form of Medical Certificate of Cause Of Death

4.3 MULTIPLE PRIMARIES

The cancer registry counts tumours not persons. Cancer patients may develop independent cancers in their lifetime. Before registering a case as a new tumour, consider:

- Is the lesion an extension, or metastasis of an existing tumour?
- Is it a recurrence[2] of an earlier tumour?

If the response to the above questions is "NO", it should be considered a new primary and a separate registration/notification form should be prepared, including morphology, behaviour, basis of diagnosis etc.

When the data are entered into CanReg, the user will be asked to confirm whether the tumour is a new primary or an extension or recurrence of an existing cancer. The rules of the International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR) are used by CanReg (IARC Internal Report No.2004/02, http://www.iacr.com.fr/images/doc/MPrules_july2004.pdf). These rules were developed for international comparisons when reporting cancer incidence and survival. They are reproduced in detail in Appendix 1.

USE ONE CANCER REGISTRATION/NOTIFICATION FORM FOR EACH PRIMARY TUMOUR DIAGNOSED.

[2] When cancer returns after a period of remission, it is considered a "recurrence". A cancer recurrence happens because, in spite of the best efforts to treat or clear off cancer, some cells remain. These cells could be in the same place where the first cancer originated, or they could be in another part of the patient's body. These cancer cells may have been dormant for a period of time, but eventually they continued to multiply, resulting in the reappearance of the cancer.

4.4 THE VARIABLES TO BE RECORDED ON EACH CASE

Mandatory variables (must be abstracted) in red

4.4.1 CANCER REGISTRY NUMBER (CRN)

A unique number assigned by the registry to each patient. This number is written on all documents and items of information relating to the patient. The first four digits of the CRN are usually the year when the patient was registered.

Example: 2015- 0001 is the CRN assigned to the first patient registered in 2015.

[CanReg-5 automatically allocates a "patient number" to all records of the same individual. It does not have to be recorded on the registration form]

4.4.2 PATIENT INFORMATION

ID number

Record the personal identification number (the national identity number, social security number) which is unique to the individual, whenever it can be found.
Abstract in detail the complete number, including any check digits when they exist.

Names

Whenever possible:

- Give the full names of the patient.
- They should be recorded as first name/personal name.
- Followed by family name.

For married women who have taken the name of their husband:

- The family name of the husband should be used.
- The patient's maiden name (unmarried name or name at birth) should be indicated under the heading 'Maiden name"

Titles such as Dr, Reverend, El Haj, etc. should be

- Entered on AFTER the first name, like this: Beatrice (Sister); Peter (Prof.) ….

Date of birth

When present in the record, enter as day, month and year (dd/mm/yyyy). If any part of this information is not known, record as unknown or not specified (e.g. 99/99/2014).

Age

This refers to the age in years at the incidence date (see below). It MUST be recorded, as the patient's age on his/her last birthday; do not round off to the next birthday.

- If the birthdate is known, check whether the given age is correct or not.
- Infants aged less than 12 months of age, record as 00.
- For persons aged 98 or MORE, record as 98.
- If not possible to find the age, enter 99 (age unknown).

Sex

Enter as 1 for male, 2 for female.
If the sex is not recorded, this may be inferred from the given name[3] and from the wording of the hospital summary. In very rare instances, the sex cannot be determined or there may have been a sex change. This information should be recorded.

[3] The fact that some names are unisex should be taken into consideration.

Usual residence address

Record as much detail as possible of the patient's usual residence. Ideally this should include the number, street, city or municipality, province and country residence. The usual residence is where the patient would be counted, if a census took place.

It **MUST** be distinguished from the patient's temporary address at the time of admission, for example a patient from the country may come to the city for medical treatment and stay temporarily with friends or relatives. His/her address in the country is the permanent address and the address in the city is the temporary address. Record the person's country of birth if this variable is required.

Telephone number

Record all the telephone numbers (fixed line, mobile phones) of the patient **AND** of the next of kin.

Ethnic group

Indicate to which ethnic group (tribe, or language group) the patient belongs.
There may be some problems in classifying individuals of mixed heritage. Record all the details. When abbreviations are used in the medical record, be sure to know exactly what the abbreviations mean.

4.4.3 INFORMATION ON THE TUMOUR

Date of incidence

The date of the first event (of the six listed below) to occur chronologically should be chosen as incidence date. If an event of higher priority occurs within three months of the date initially chosen, the date of the higher priority event should take precedence. Order of declining priority:

1. Date of first histological or cytological confirmation of this malignancy (with the exception of histology or cytology at autopsy). This date should be, in the following order:
    a date when the specimen was taken (biopsy)
    b. date of receipt by the pathologist
    c. date of the pathology report
2. Date of admission to the hospital because of this malignancy.
3. When evaluated at an outpatient clinic only: date of first consultation at the outpatient clinic because of this malignancy.
4. Date of diagnosis, other than 1, 2 or 3. It may be date of first clinical investigation procedure fort the malignancy e.g.: MRI reports, CT scan reports etc.
5. Date of death, if no information is available other than the fact that the patient has died because of a malignancy.
6. Date of death, if the malignancy is discovered at autopsy.

Whichever date is selected, the date of incidence should **NOT** be later than the date of the start of the treatment, or decision not to treat, or date of death.

Basis of diagnosis

The medical records should be studied carefully to determine the different methods used to confirm the diagnosis of cancer. The most valid basis of diagnosis or the most conclusive method of confirmation should be noted down on the abstract. If additional information becomes available later, the most valid basis for diagnosis should be updated.

The suggested codes (Table 4.1) are hierarchical, so that the higher number represents the more valid basis, and should thus be used for this purpose. If there is no information on how the diagnosis had been made (information obtained from an automated source, for example) the code 9 (Unknown) should be used.

Code 0 is used for Death Certificate Only – that is, cases registered for which the only available information on cancer was on a death certificate, and where follow-back attempts have been unsuccessful. This category does not include cases first coming to the registry's attention by means of a death certificate mentioning cancer for which other bases of diagnosis became available.

Code 6 should be used when a histology examination has shown cancer to be present – but the specimen examined contained a metastasis, and was not from the site of origin (primary site) of the tumour. This is often the case when the specimen is a lymph node.

CODE	DESCRIPTION	CRITERIA
0	**Death Certificate Only**	The only information to the registry is from a death certificate.
Non Microscopic		
1	**Clinical**	Diagnosis made before death, but without the benefit of any of the following (2-7)
2	**Clinical investigation**	To include all diagnostic techniques, including x-ray, endoscopy, imaging, ultrasound, exploratory surgery (e.g., laparotomy) and autopsy, without a tissue diagnosis.
4	**Specific tumour markers**	To include biochemical and/or immunological markers which are specific for a tumour site (Table 4.2).
Microscopic		
5	Cytology	Examination of cells whether from a primary or secondary site, including fluids aspirated using endoscopes or needles. Also to include the microscopic examination of peripheral blood films and trephine bone marrow aspirates.
6	Histology of a metastasis	Histological examination of tissue from a metastasis, including autopsy specimens.
7	Histology of a primary tumour	Histological examination of tissue from the primary tumour, however obtained, including all cutting techniques and bone marrow biopsies. Also to include autopsy specimens of a primary tumour.
9	Unknown	

Table 4. 1 IARC - IACR Basis of Diagnosis Codes

Specific tumour markers	
Human Chorionic Gonadotrophin (HCG)	In diagnosis of choriocarcinoma (usually >100,000 iu in urine)
Prostate Specific Antigen (PSA)	In diagnosis of prostate carcinoma (usually >10 µg/l serum)
Alphafetoprotein (AFP)	In diagnosis of hepatocellular carcinoma (usually >200 ng/ml serum)
Catecholamine degradation products (HVA, VMA)	In diagnosis of neuroblastoma
Elevated serum immunoglobulins	Myeloma (IgG >35g/l or IgA > 20g/l) Waldenström's macroglobulinaemia (IgM > 10g/l)
Urinary immunoglobulins	Myeloma (light chain excretion > 1g/24hr)

Table 4. 2 Specific Tumour Markers

Primary site

Carefully review all reports contained in the clinical record and record the site in which the tumour originated. The primary site may at times be determined by a pathologist reviewing tissue from a secondary site (e.g. a primary carcinoma of lung diagnosed by excision and microscopic review of lymph nodes). It is also possible to deduce a primary site from the determination of a specific morphology (e.g. a nodular melanoma of the neck indicates a malignancy of the skin of the neck). [See RULE H (site associated morphology), in the ICD-O coding section, page 29].

Sites such as 'head', 'thorax', 'limb', 'pelvis', and ‘abdomen’ are poor descriptors of site, since a tumour may arise in a number of tissues (skin, soft tissue and bone) within these sites. It is important to extract all the diagnostic information available in the record.

If there is no mention of the primary site in the record, but a secondary site(s) has been identified, note all available information regarding the secondary site(s) – **BUT DO NOT CODE OR ENTER THE SECONDARY SITE INTO THE COMPUTER**. The information on the primary site may be added at a later date if it becomes available.

Morphology

In abstracting histology, record the complete histological diagnosis as stated in the pathology report's Final Diagnosis. Do not modify the pathologist's final diagnosis by picking up descriptive terms found in the microscopic description of the tissue.

If conflicting statements exist regarding the diagnosis, prefer statements from the pathology reports over other statements.

If the histological diagnosis is stated using only non-specific terms such as 'malignant neoplasm’, ‘cancer' or ‘malignant tumour', abstract these terms until more detailed information becomes available.

Behaviour

The behaviour of a tumour is the way it acts within the body. The behaviour of the tumour is coded as the 5th digit of the morphology code (after the "/"). Table 4.3 shows the spectrum of behaviours. A tumour can grow in place without the potential for spread (/0, benign); it can be malignant but still growing in place (/2, non-invasive or in situ); it can invade surrounding tissues (/3, malignant, primary site).

Behaviour codes /6, malignant, metastatic site, and /9, malignant, uncertain whether primary or metastatic site, must not be used.

Always enter the PRIMARY site (with /3 to indicate a malignant tumour). If the site of the primary cancer is unknown, this should be noted and the appropriate ICD code will be given C80.9 (unknown primary site).

Code	
/0	Benign
/1	Uncertain whether benign or malignant Borderline malignancy Low malignant potential Uncertain malignant potential
/2	Carcinoma in situe Intraepithelial Noninfiltrating Noninvasive
/3	Malignant, primary site
/6*	Malignant, metastatic site Malignant, secondary site
/9*	Malignant, uncertain whether primary or metastatic site

* Not used by cancer registries

Table 4. 3 Behaviour codes (ICD O-3)

Laterality

This should be recorded for all paired organs, but as a minimum for lung, breast, eye, ovary, testis and kidney.

Stage

FOR ADULTS: record **stage of disease** as it is found in the case record.
If it is present, record the staging system that was used:-

- UICC/AJCC stage is the most widely used (and if the staging system is not specified, assume it is this one)
- FIGO - Female reproductive site cancers was developed by the International Federation of Gynaecology and Obstetrics.
- DUKE's - The Duke's staging system is a classification system for colorectal cancers.

Unless you have been trained and are authorised to do so, **DO NOT** assign stage to a cancer if it is not noted in the patient's medical records.

FOR CHILDREN (age less than 15): Stage should be recorded according to the "Toronto" system, as described in Childhood Cancer Staging Rules for AFCRN Registries.

TNM

The extent of disease should be recorded in terms of the three digit code of the TNM system. The rules for coding according to the TNM system are described in TNM Classification of Malignant Tumours, 8th Edition, 2017 (Brierley, Gospodarowicz and Wittekind).

The TNM system is not used for the coding of the extent of lymphomas, leukaemias, brain tumours and childhood cancers (defined as < 15 years of age at diagnosis).

Staging should be done at the time of initial diagnosis. It is based on information that can be either clinical (c), which is the stage before any treatment, or pathological (p), which is the post-surgical histo-pathological classification.

In the absence of surgery, staging is based upon examinations carried out prior to medical treatment, or radiotherapy, or during the hospital stay when these treatments were started, or a decision made to withhold them.

The detection of metastatic disease **after** the first course of treatment (including during adjuvant treatment or hormonal therapy) does **NOT** change coding of extent of disease at diagnosis.

pTNM vs. cTNM

When the stage/extent of the cancer is recorded in the clinical and/or pathological records according to the TNM system, these codes should be registered.

Record stage from pathology - pT (rather than cT) and pN (rather than cN), if they are available.

When T, and/or N, and/or M have not been explicitly recorded in the clinical/pathological records, the cancer registrar should attempt to score extent of disease according to the Essential TNM scheme (see section 5.7)

4.4.4 TREATMENT

Record any treatment described in the patients' records initiated within 4 months from incidence date. This includes therapy given at the reporting hospital as well as those given in other facilities. Treatment is considered as a specific therapy which controls or destroys cancer tissues both at the primary and metastatic sites. **Cancer-directed treatments** include surgery, radiotherapy, chemotherapy, hormonal therapy, immunotherapy and palliative care. Also record any other care that the patient received.

Record the **DATE** on which each of these treatments was started.

Surgery

It is total or partial surgical removal of the tissue of the primary or the metastatic sites. It is performed most of the time after diagnosis.

Radiotherapy

Include external or beam radiotherapy, or internal radiation.

External radiotherapy uses X-rays from cobalt or linear accelerator machines, electrons, and more rarely other particles such as protons to destroy cancer cells in the treated area by damaging the DNA within these cells.

Internal radiation A source of radiation is put inside the body. One form of internal radiation therapy is called brachytherapy**,** where the radiation source is a solid in the form of seeds, ribbons, or capsules, which are placed in the body or near the cancer cells. Used for cancers of the head, neck, breast, uterus, cervix, prostate, gall bladder, oesophagus, eye, and lung. Internal radiation can also be in a liquid form used with people who have thyroid cancer or non-Hodgkin's lymphoma.

Chemotherapy

Chemotherapy is usually given as an intravenous injection or drip, but sometimes drugs are in the form of a tablet. A list of chemotherapeutic drugs is given in Appendix 2.

Hormonal therapy

Hormonal therapy medicines are whole-body (systemic) treatment for hormone-receptor-positive cancers, such as some breast and prostate cancers. They include:

Hormonal Agent	Brand Name(S)	Hormonal Agent	Brand Name(S)
Anastrozole	Arimidex	Goserelin (Breast)	Zoladex
Abiraterone acetate	Zytiga	Goserelin (Prostate)	Zoladex, Zoladex LA, Novgos
Bicalutamide	Casodex	Letrozole	Femara
Buserelin	Suprefact	Leuprorelin acetate	Prostap SR, Prostap 3
Cyproterone	Cyprostat	Medroxyprogesterone	Provera
Degarelix	Firmagon	Megestrol acetate	Megace
Diethylstilbestrol	Stilboestrol	Tamoxifen	Nolvadex, Tamoxen, Tamosin, Tamofen
Exemestane	Aromasin	Toremifene	Fareston
Flutamide	Drogenil	Triptorelin	Decapeptyl SR, Gonapeptyl Depo
Fulvestrant	Faslodex		

Immunotherapy

Immunotherapy (also called biologic therapy or biotherapy) uses materials either made by the body or in a laboratory to improve, target, or restore immune system function. There are several types of immunotherapy, including non-specific immunotherapies (e.g. aldesleukin, interferon) and monoclonal antibodies.

Immunotherapy Agent	Brand Name(S)	Immunotherapy Agent	Brand Name(S)
Aldesleukin	Proleukin	Iodine-131 tositumomab	Bexxar
Interferon	IntronA, Roferon-A	Ipilimumab	Yervoy
90Y-Ibritumomab tiuxetan	Zevalin	Panitumumab	Vectibix
Bevacizumab	Avastin	Rituximab	Mabthera
Cetuximab	Erbitux	Trastuzumab	Herceptin
Gemtuzumab	Mylotarg		

4.4.5 SOURCE OF INFORMATION

It is important to record the source of information every time a cancer case is identified from one of the sources. The source may be a hospital, clinic, hospice, laboratory, or a death certificate.

Clearly write the details of the source (ward or service of a hospital, which laboratory etc.) so that it can be coded.

Record the patient's FILE NUMBER as provided on the cover of the medical records file or the laboratory reference number (e.g. Pathology number) from the report.

For each source, record the DATE:

- ☞ For hospital cases, the date of admission to the hospital
- ☞ For out-patients – the date of consultation
- ☞ For laboratories, the date of examination (as given in the laboratory/ X-Ray report).

This is very important – these numbers will be needed if the case record is needed to be traced using the cancer registry database.

4.4.6 FOLLOW UP

It is important to have follow-up information of each cancer patient registered.

Follow-up information is important when estimating cancer survival as a measure of outcome. Information expected is either the patient is alive or dead or unknown (lost to follow-up). Follow-up procedures employed by the registry should be clearly specified.

- Active follow up may be done for special studies - by contacting the patient's physician or the patients themselves (by telephone, mail, or home visits).
- Access to death register/death certificates allows them to be used as a passive method of follow up.

Date of last contact

Refers to the latest date for which there is ANY information about the patient. It may be:

- The date he/she was last known to be alive
- The date of death

At the time of the first registration, this will probably be the date of discharge from hospital (or of outpatient appointment).

More information (later dates) may come from follow-up visits to the same hospital, or from admission elsewhere (e.g. radiotherapy or hospice care).

For death certificate registrations, Date of Last Contact = Date of Death.

Give date of last contact as the complete date, including day, month and year.

Status at last contact

Record whether the patient was alive or dead on the Date of Last Contact.
Codes: 1 Alive 2 Dead

Try **NOT** to code 9 ("Unknown") – the status of the patient on the date of last follow up should never be "unknown" when information on them was traced!

Cause of death

If the patient was alive on date of last contact, enter "Not applicable" (Code = 8) in CanReg.

If the patient was dead, there are two options for recording the cause of death:

- 1 Dead of cancer 2 Dead of other cause 9 Not known.

OR

- The underlying cause of death as specified in the death certificate.

Where is the "underlying cause of death"? If the death certificate uses the WHO recommended standard form of medical certification (Fig 4. 2, page 11):

In Part I, the cause leading directly to death is reported on line (**a**),

The intervening antecedent condition (if any) on line (**b**), and

The underlying cause of death on line (**c**).

If the entry on line (**a**), or on lines (**a**) and (**b**)**,** completely describe the sequence of events leading to death, then it is no longer necessary to put an entry on line (**c**).

Part II is for any condition which may contribute to death but is not related to the disease or condition causing the death.

For coding the underlying cause of death, use the appropriate codes of the International Classification of Diseases (ICD-10).

5. CODING

Abstracted information should be coded first before captured into the computer.

5.1 PLACE OF RESIDENCE

Place of residence coding should follow the administrative subdivisions used in the census (for which population numbers are available), and, ideally, the same coding system. The coding scheme should be hierarchical – for example, going from region – district- sub-district – village.

[Example- Cancer registries of Mozambique]

HASC	CG	
MZ.NS.--	100	**Niassa district**
MZ.CD.--	200	**Cabo Delgado district**
MZ.NM.--	300	**Nampula district**
MZ.ZA.--	400	**Zambezia district**
MZ.TE.--	500	**Tete district**
MZ.MN.--	600	**Manica district**
MZ.SO.--	700	**Sofala district** desconhecido
MZ.SO.BC	→→	Cidade de Beira →→→→→
MZ.SO.BU	702	Buzi
MZ.SO.CA	703	Caia
MZ.SO.CM	704	Chemba
MZ.SO.CR	705	Cheringoma
MZ.SO.CB	706	Chibabava
MZ.SO.DO	707	Dondo
MZ.SO.GO	708	Gorongosa
MZ.SO.MC	709	Machanga
MZ.SO.MG	710	Maringue
MZ.SO.MM	711	Marromeu
MZ.SO.MU	712	Muanza
MZ.SO.NH	713	Nhamatanda
MZ.IN.--	800	**Inhambane district**
MZ.GA.--	900	**Gaza: district**
MZ.MP.--	*000*	**Maputo province**
MZ.MC.--	*010*	**Maputo city**: district desconhecido
MZ.MC.DU	*011*	Distrito Urbano Nº 1
MZ.MC.DD	*012*	Distrito Urbano Nº 2
MZ.MC.DT	*013*	Distrito Urbano Nº 3
MZ.MC.DQ	*014*	Distrito Urbano Nº 4
MZ.MC.DC	*015*	Distrito Urbano Nº 5
MZ.--.--	*080*	Mocambique: prov. desconhecido
	090	Outros países: África
	091	Outros países: Europa
	092	Outros países: Ásia
	093	Outros países: América
	099	Desconhecido

Cidade de Beira			
720	Barrio desconhecido		
721	Barrio 1	*733*	Barrio 13
722	Barrio 2	*734*	Barrio 14
723	Barrio 3	*735*	Barrio 15
724	Barrio 4	*736*	Barrio 16
725	Barrio 5	*737*	Barrio 17
726	Barrio 6	*738*	Barrio 18
727	Barrio 7	*739*	Barrio 19
728	Barrio 8	*740*	Barrio 20
729	Barrio 9	*741*	Barrio 21
730	Barrio 10	*742*	Barrio 22
731	Barrio 11	*743*	Barrio 23
732	Barrio 12	*744*	Barrio 24

5.2 ETHNIC GROUP

[Example – Eastern Cape Registry, South Africa]

1 WHITE
2 ASIAN
3 COLORED
4 BLACK: XHOSA
5 BLACK: ZULU
6 BLACK: SOTHO
7 BLACK: TSWANA
8 BLACK: OTHER
9 UNKNOWN

5.3 BASIS OF DIAGNOSIS

Code basis of diagnosis according to the codes on the right.

This coding scheme permits the distinction between tumours diagnosed on the basis of histology of a metastasis, or from the primary site, making the use of behaviour code /6 (and /9) unnecessary.

Code

0. Death certificate only

Non-microscopic

1. Clinical

2. Clinical investigation

4. Specific tumour markers

Microscopic

5. Cytology or haematology

6. Histology of a metastasis

7. Histology of a primary tumour

9. Unknown

5.4 SITE AND MORPHOLOGY OF TUMOUR (ICD-O)

Coding is according to the International Classification of Diseases for on Oncology (3rd Edition). Full instructions for coding are given in that book.

Topography

Code the ICD-O code (ranging from C00.0 to C 80.9) corresponding to the primary site of origin of the tumour. The decimal point (.) indicates subdivisions or sub-sites of the three character categories.

Do **NOT** code the site of any secondary/metastatic cancer (which may have been examined, for example, by a pathologist).

Code C80.9 is for unknown primary site.

Morphology

The five digit numerical list ranges from 8000/0 to 9989/1. The first four digits indicate the specific histology and the fifth digit after the slash is the behaviour of the tumour.

The diagnosis may be stated using non-specific terms instead of a specific histological type; for example malignant neoplasm, cancer etc. This will normally be the case if the basis of diagnosis is non-microscopic.

In such cases, morphology code 8000/3 should be used.

It is very unlikely (or impossible) for most specific morphological diagnoses to have been made without a histological (or cytological) examination. However, certain combinations are exceptions to this general rule, as shown in Table 5. 1.

MORPHOLOGY		Most Valid	Other criteria
Code	Description	Basis	
8800	(Sarcoma NOS)	2	
9590	Lymphoma NOS	1 or 2	
9800	Leukaemia NOS	1 or 2	
8720	Melanoma	1 or 2	
9140	Kaposi's sarcoma	1 or 2	HIV positive (exc. Africa)
8960	Nephroblastoma	2	Age 0-8
9100	Choriocarcinoma	4	Female, and age 15-49
9500	Neuroblastoma	2 or 4	Age 0-9
9510	Retinoblastoma	2	Age 0-5
9732	Myeloma	4	Age 40+
9761	Waldenström's macroglobulinaemia	4	Age 50+
8170	Hepatocellular carcinoma	4	
8150-8154	Islet cell tumours, gastrinomas	4	
9380	Glioma	2	C71.7 (brain stem)
9384/1	Subependymal giant cell astrocytoma	2	Tuberous sclerosis patient
9530-9539	Meningioma	2	C70
9350	Craniopharyngioma	2	
8270-8281	Pituitary tumours	4	C75.1

Table 5. 1 Combinations of specific morphology codes, and non-microscopic basis of diagnosis codes, which are considered acceptable

"NOS" means "Not Otherwise Specified". In the numerical list and in the alphabetical indexes,"NOS" is used to indicate that other modifiers of the term are listed elsewhere.

SUMMARY OF PRINCIPAL RULES FOR USING ICD-O, THIRD EDITION

RULE A. ***Topographic regions and ill-defined sites:*** If the diagnosis does not specify the tissue of origin, code the appropriate tissues suggested in the alphabetic index for each ill-defined site in preference to the "NOS" category. Ill-defined sites, such as "arm", have several component tissues. For example, "squamous cell carcinoma of the arm" should be coded to C44.6 (skin of arm) rather than to C76.4 (arm, NOS).
See *ICD-O 3 Manual: Coding Guidelines*, *page 24*. There are a few exceptions to this, such as chin and forehead, because these regions are predominantly composed of skin, and the NOS category was therefore assigned to skin.

RULE B. ***Prefixes:*** If a topographic site is modified by a prefix such as peri-, para-, or the like which is not specifically listed in ICD-O, code to the appropriate ill-defined subcategory C76 (ill-defined site), unless the type of tumour indicates origin from a particular tissue. This general rule also applies to imprecise phrases such as "area of' or "region of '.
See *ICD-O 3 Manual: Coding Guidelines*, *page 25*.

RULE C. ***Tumours involving more than one topographic category or subcategory:*** Use subcategory".8" when a tumour overlaps the boundaries of two or more categories or subcategories and its point of origin cannot be determined.
See *ICD-O 3 Manual: Coding Guidelines*, page 25, and *Note, page 45*.

RULE D. ***Topography codes for lymphomas:*** If a lymphoma involves multiple lymph node regions, code to C77.8 (lymph nodes of multiple regions). Code extra nodal lymphomas to the site of origin, which may not be the site of the biopsy. If no site is indicated for a lymphoma, code to C77.9 (lymph node, NOS). Lymphomas occur in specific sites, for example stomach, as well as in one or more lymph nodes and therefore are not assigned a site-specific topography code. Lymphomas occurring in specific sites are called extra nodal.
See *ICD-O 3 Manual: Coding Guidelines*, *page 26* and *the malignant lymphoma section, page 13*.

RULE E. ***Topography code for leukaemias:*** Code all leukaemias except myeloid sarcoma (M-9930/3) to C42.1 (bone marrow).
See *ICD-O 3 Manual: Coding Guidelines, page 26*.

RULE F. ***Behaviour code in morphology:*** Use the appropriate 5th digit behaviour code (table below) even if the exact term is not listed in ICD-O. The use of the 5th digit behaviour code is explained in the ICD-O 3 Manual: Coding Guidelines.

Code	Definition
0	**Benign**
1	**Uncertain (Benign/Malignant)**
2	**In-situ**
3	**Malignant**

The appropriate 5th digit code should be used even if the exact term is not listed in ICD-O; for example, "benign chordoma" as a diagnosis should be coded M-9370/0. If the pathologist states that the behaviour differs from the usual behaviour as given in ICD-O, code as the pathologist indicates.

RULE G. ***Grading or differentiation code:*** **Assign the highest grade or differentiation code described in the diagnostic statement.** The use of the 6th digit for grading or differentiation of solid tumours is explained in the

ICD-O 3 Manual: Coding Guidelines, page 30. If a diagnosis indicates two different degrees of grade or differentiation (such as "well and poorly differentiated" or "grades 11-m"), code to the higher grade.

This 6th digit may also be used for identifying the cell origin for lymphomas and leukaemias. In these lymphatic and hematopoietic diseases, T-cell (code 5), B-cell (code 6), Null cell (code 7), and NK cell (code 8) take priority over grade codes 1 to 4.

RULE H. *Site-associated morphology terms:* Use the topography code provided when a topographic site is not stated in the diagnosis. This topography code should be disregarded if the tumour is known to arise at another site. The appropriate site-specific codes are listed in parentheses after morphology terms for neoplasms that usually occur in the same site or tissue, for example "retinoblastoma" (C69.2). If no site is indicated in the diagnosis, use the suggested code.

If the site given differs from the site-specific code indicated for the morphologic type, use the appropriate code for the site given. This should be done only after thoroughly reviewing the case to ascertain that the neoplasm at the site mentioned is not a metastasis. Only three-character codes are given for some sites, for example C44.- (skin), because the appropriate fourth-digit cannot be assigned in advance.
See ICD-O 3 Manual: Coding Guidelines, page 32.

Certain neoplasms have names that could be interpreted as implying a topographic location (pseudotopographic morphology terms), but these entities should not necessarily be coded to that site. For example, bile duct carcinoma is a tumour frequently arising in intrahepatic bile duct of liver (C22.1).
See ICD-O 3 Manual: Coding Guidelines, page 33.

RULE J. *Compound morphology diagnoses:* Change the order of word roots in a compound term if the term is not listed in ICD-O. Not all forms of compound words are listed. For example,
"myxofibrosarcoma" is not in ICD-O but "fibromyxosarcoma" is. Check various permutations of the word roots if the first term is not found.
See ICD-O 3 Manual: Coding Guidelines, page 33.

RULE K. *Coding multiple morphology terms:* When no single code includes all diagnostic terms, use the numerically higher code number if the diagnosis of a single tumour includes two modifying adjectives with different code numbers. If a term has two or more modifying adjectives with different code numbers, code to the one with the highest code number, as it is usually more specific. For example a tumour described as a “transitional cell epidermoid carcinoma” should not be reported twice (one as transitional cell carcinoma M8120/3 and the other as “epidermoid carcinoma”M8070/3. Code 8120/3 should be assigned since it is the highest.
See ICD-O 3 Manual: Coding Guidelines, page 34.

5.5 CODING OF PRE-INVASIVE CERVICAL LESIONS

According to the WHO, a squamous intraepithelial lesion that presents a significant risk of developing into invasive cancer if left untreated has the following synonyms:

- High-grade squamous intraepithelial lesion (HSIL)
- Cervical intraepithelial neoplasia grade 2 (CIN 2);
- Cervical intraepithelial neoplasia grade 3 (CIN 3);
- Moderate squamous dysplasia;
- Severe squamous dysplasia;
- Squamous carcinoma in situ (CIS).

All are coded in the ICD-O as 8077/2

5.6 CODING METASTATIC CANCERS

Adapted from ICD-10 coding rules (WHO ,2010)[4]

The expression "metastatic" is a problem mainly in the English language.
Neoplasms qualified as metastatic are always malignant, either primary or secondary.
However, the adjective "metastatic" is used in two ways, sometimes meaning a secondary from a primary elsewhere and sometimes denoting a primary that has given rise to metastases.

Although malignant cells can metastasize anywhere in the body, certain sites are more common than others and must be treated differently. These sites are listed in Table 5.2 below

Table 5.2 Common sites of metastases

Bone	Mediastinum
Brain	Meninges
Diaphragm	Peritoneum
Ill-defined sites (sites classifiable to C76)	Pleura
Liver	Retroperitoneum
Lung (see special instructions at (f))	Spinal cord
Lymph nodes	

(a) Malignant neoplasm "metastatic from"

If a malignant neoplasm is described as "metastatic from" a specified site, that site should be considered primary.

Example 1: Metastatic teratoma from ovary

The expression "metastatic teratoma from ovary" implies that the neoplasm originated in the ovary. Code to ovary (C56).
This also applies to sites on the list of common sites of metastases.

Example 2: Metastatic mesothelioma from peritoneum

[4] International statistical classification of diseases and related health problems. - 10th revision, edition 2010.
Volume 2. Instruction manual

A “metastatic mesothelioma from peritoneum” is primary in the peritoneum, although peritoneum is one of the sites listed in Table 3.
Code to malignant mesothelioma of peritoneum (C45.1).

(b) Malignant neoplasm “metastatic to”

A malignant neoplasm described as “metastatic to” a specified site should be interpreted as a secondary neoplasm of the specified site, whether the site is on the list of common sites of metastases or not. Code to malignant neoplasm of unknown primary site (C80.9) if no primary site is indicated.

Example 3: Metastatic carcinoma to the rectum

The expression “metastatic to” indicates that rectum is a secondary site.
Code malignant neoplasm of unknown primary site (C80.9), since no primary site is indicated.

If the morphology code has a “preferred site” in ICD-O (see Rule H, section 5.6) use the topography code provided as the primary site, when this is not stated in the diagnosis

Example 4: Metastatic osteosarcoma to brain

The expression “metastatic to brain” indicates that brain is a secondary site. However, the osteosarcoma is indexed to malignant neoplasm of bone (C40._ ; C41._) in the alphabetical Index of ICD-O.
Code unspecified malignant neoplasm of bone (C41.9)

(c) Malignant neoplasm metastatic of site A to site B

A malignant neoplasm described as metastatic of site A to site B should be interpreted as primary of site A and secondary of site B.

Example 5: Metastatic cancer of liver to brain

The expression “metastatic of liver to brain” indicates that the malignancy originated in the liver and spread to the brain.
Code to primary cancer of liver (C22.9).

(d) “Metastatic” malignant neoplasm on the list of common sites of metastases

A “metastatic” neoplasm is considered secondary if the site is on the list of common sites of metastases.
A neoplasm of a site in Table 5.2 is considered secondary, even if no other neoplasm is mentioned in the report. Note that a secondary malignant neoplasm should not be selected as the Primary site.
If no primary tumour is reported, code the case to malignant neoplasm of unspecified site (C80.9).

Example 6: Metastatic brain cancer

Brain is one of the sites in Table 5.1, and the “metastatic” brain cancer is considered secondary. There is no primary neoplasm reported.
Therefore, code to malignant neoplasm of unknown primary site (C80.9).

(e) “Metastatic” malignant neoplasm not on the list of common sites of metastases

If a site that is not on the list of common sites of metastases is qualified as “metastatic” or “metastatic of”, consider it primary and code to malignant primary of that particular site.

Example 7: Cervix cancer, metastatic
Cervix is not in Table 3, and the "metastatic" cervix cancer is therefore considered primary.
Code to malignant neoplasm of cervix (C53.9).

(f) "Metastatic" cancer of lung

The lung poses special problems in that it is a common site for both metastases and primary malignant neoplasms. Lung cancers can be both a primary or secondary, depending on other neoplasms reported in the source documents, if any.
If the only malignancy mentioned is "metastatic" neoplasm of lung, *code to primary malignant neoplasm of lung.*

Example 8: Metastatic carcinoma of lung
Code to primary malignant neoplasm of lung (C34.9) since no other site is mentioned.

If another malignancy is mentioned that is not on the list of common sites of metastases, consider lung secondary.

Example 9: Metastatic cancer of lung and stomach cancer
Since stomach cancer is also mentioned, "metastatic cancer of lung" is considered secondary. Select and code stomach cancer (C16.9) as the primary site.

(g) "Metastatic" neoplasm of a specific morphology

If the morphological type has a "preferred site" in ICD-O (see Rule H, section 5.6) and the site reported in the source documents indicates the same type of tissue, use the topography code provided as the primary site.

Example 10: Metastatic osteosarcoma of femur
Code to malignant neoplasm of long bones of lower limb (C40.2).

If the morphological type has a preferred site (Rule H of ICD-O) and the site reported indicates a *different* type of tissue, code to the unspecified site for the morphological type.

Example 11: Metastatic nephroblastoma of hilar lymph nodes
Code to unspecified site for kidney (C64.9).

5.7 CODING TNM AND STAGE

5.7.1 Coding of STAGE

ADULTS

Table 5.3 Codes for Stage in adults

CODE	UICC/AJCC STAGE	FIGO	Hodgkin Lymphoma
1	Stage I	Stage I	Stage I
1A	Stage IA	Stage IA	Stage IA
1B	Stage IB	Stage IB	Stage IB
2	Stage II	Stage II	Stage II
2A	Stage IIA	Stage IIA	Stage IIA
2B	Stage IIB	Stage IIB	Stage IIB
2C	Stage IIC		
3	Stage III	Stage III	Stage III
3A	Stage IIIA	Stage IIIA	Stage IIIA
3B	Stage IIIB	Stage IIIB	Stage IIIB
3C	Stage IIIC	Stage IIIC	
4	Stage IV	Stage IV	Stage IV
4A	Stage IVA	Stage IVA	Stage IVA
4B	Stage IVB	Stage IVB	Stage IVB

CHILDREN

Table 5.4 Codes for Stage in children

<table>
<tr><th>CODE</th><th>STAGE</th><th colspan="2">CANCERS</th></tr>
<tr><td>C+</td><td>CNS+</td><td colspan="2" rowspan="2">Acute lymphoblastic leukaemia</td></tr>
<tr><td>C-</td><td>CNS-</td></tr>
<tr><td>Lx</td><td>Localised</td><td rowspan="2">Soft tissue sarcomas
Wilms tumour (nephroblastoma)
Bone cancers Neuroblastoma
Liver cancers
CNS/Brain cancers</td><td rowspan="3">Retinoblastoma
Testicular cancer
Ovarian cancer</td></tr>
<tr><td>Mx</td><td>Metastatic</td></tr>
<tr><td>Rx</td><td colspan="2">Regional</td></tr>
<tr><td>LR</td><td>Locoregional</td><td colspan="2" rowspan="2">Neuroblastoma</td></tr>
<tr><td>MS</td><td>Metastatic - limited</td></tr>
<tr><td>Lx</td><td>Limited</td><td colspan="2" rowspan="2">Non Hodgkin lymphoma,</td></tr>
<tr><td>Ax</td><td>Advanced</td></tr>
<tr><td>XX</td><td>Unknown</td><td colspan="2">ALL</td></tr>
</table>

5.7.2 Coding of TNM

- **When T, and/or N, and/or M ARE recorded in the clinical/pathological records, the cancer registrar should code the best available data (see 4.4.3).**

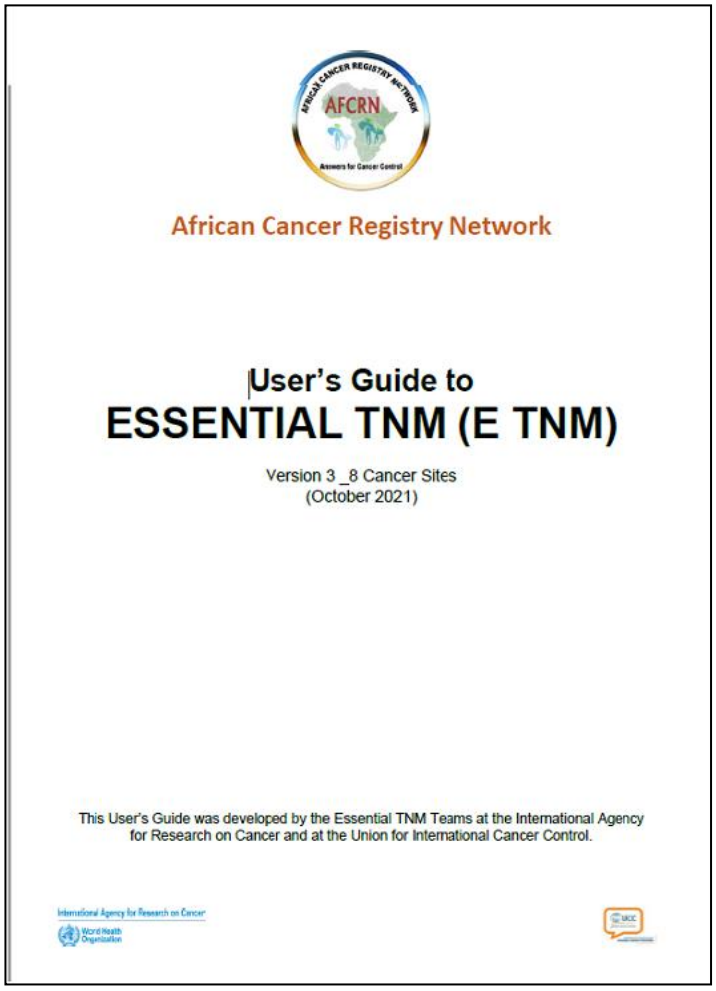
  However, if one or more of these elements are based on clinical evaluation (c TNM), and surgical/pathological information has become available at a later date, the registrar may record the appropriate Essential TNM code, if it differs from that in the record

- **When T, and/or N, and/or M have NOT been explicitly recorded in the clinical/pathological records, the cancer registrar should attempt to code extent of disease according to the Essential TNM scheme.**

  In the event of neoadjuvant therapy (i.e. systemic therapy prior to surgery) being given, information used for staging purposes should only include procedures and records prior to the initiation of this therapy.

REFER TO THE USERS GUIDE TO ESSENTIAL TNM (E-TNM)

Fig 5 Users Guide to Essential TNM (E-TNM



CODES FOR TNM

Table 5.5 Codes for TNM

	T		N		M	
	Stage	Code	Stage	Code	Stage	Code
TNM	TX	X	NX	X	MX*	X
	T1	1	N0	0	M0	0
	T1A	1A	N1	1	M1	1
	T1B	1B	N1A	1A		
	T2	2	N1B	1B		
	T2A	2A	N1C	1C		
	T2B	2B	N2	2		
	T2C	2C	N2A	2A		
	T3	3	N2B	2B		
	T3A	3A	N2C	2C		
	T3B	3B	N3	3		
	T4	4	N3A	3A		
	T4A	4A	N3B	3B		
	T4B	4B	N3C	3C		
	T4C	4C				
Essential TNM	L	LX	R-	R-	M-	M-
	L1	L1	R+	R+	M+	M+
	L2	L2	R1	R1		
	A	AX	R2	R2		
	A1	A1				
	A2	A2				

*MX, not used in TNM-8th, is included in case it is mentioned in a clinical record

5.8 SOURCE OF INFORMATION

It is a good idea to clearly separate the codes for the major types of source – for example, hospitals & clinics, laboratories & Imaging, and death certificates. This may be useful in evaluation of completeness (see section 9.2)

[Example of codes for "Source"]

HOSPITALS	
1	*University Teaching Hospital*
2	*Rigobed Government Hospital*
3	*Digby Regional Hospital*
4	*Peppa Pig Children's Hospital*
5	*Alma Mater Hospital*
6	*Hlatikhulu Hospital*
7	*Nkatama Hospital*
8	*Busy Bee Hospital*
9	*New Salop Hospital*
10	*Four Seasons Hospital*
11	*Busy Bee Clinic*
30	*Hope Hospice*
LABORATORIES	
61	*Multisystem lab*
62	*Mankayane Clinical Labs*
63	*Lancet Path Labs*
64	*Planktion diagnostic centre*
DEATH CERTIFICATE	
80	*Vital registration*
81	*UTH mortuary*

[Examples of codes for department/unit]

UNIT/DEPARTMENT	
1	*Oncology*
2	*Surgery I*
3	*Surgewry II*
4	*Urology*
5	*Gynaecology*
6	*Medicine I*
7	*Medicine II*
8	*Paediatics*
9	*ENT*
10	*Ophthalmology*
18	*Pain clinic*
30	*Out patient*

6. INPUT PROCEDURES

Notifications are received on registration/notification forms, or as computer files. Input procedures concern entering the information onto CanReg version 5 (Fig 6. 1).

Fig 6. 1 CanReg Software

The CanReg system allows input, storage, checking, back up and analysing cancer registry data. The input process also includes a number of inbuilt checks, to make sure that very obvious mistakes are flagged for correction. Incomplete of incorrect registrations cannot be CONFIRMED,and will remain in a pending state until corrected/completed.

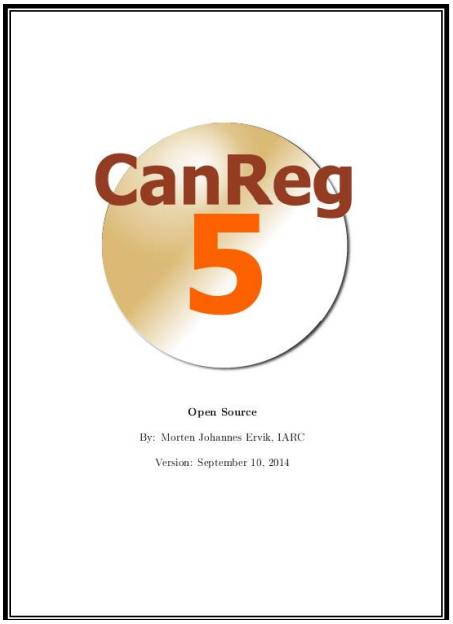


Fig 6. 2 CanReg User Guide

CanReg 5 splits the information in three tables: Patient, Tumour and Source. For each patient, you can store as many tumour records as you need, and for each tumour you can store as many source records as you need.

The CanReg manual (Fig 6. 2) gives detailed instructions on data entry procedure, including checking to see if a given patient has already a record, allowing for updating existing records, and creating new ones. Fig 6. 3 shows the basic processes involved.

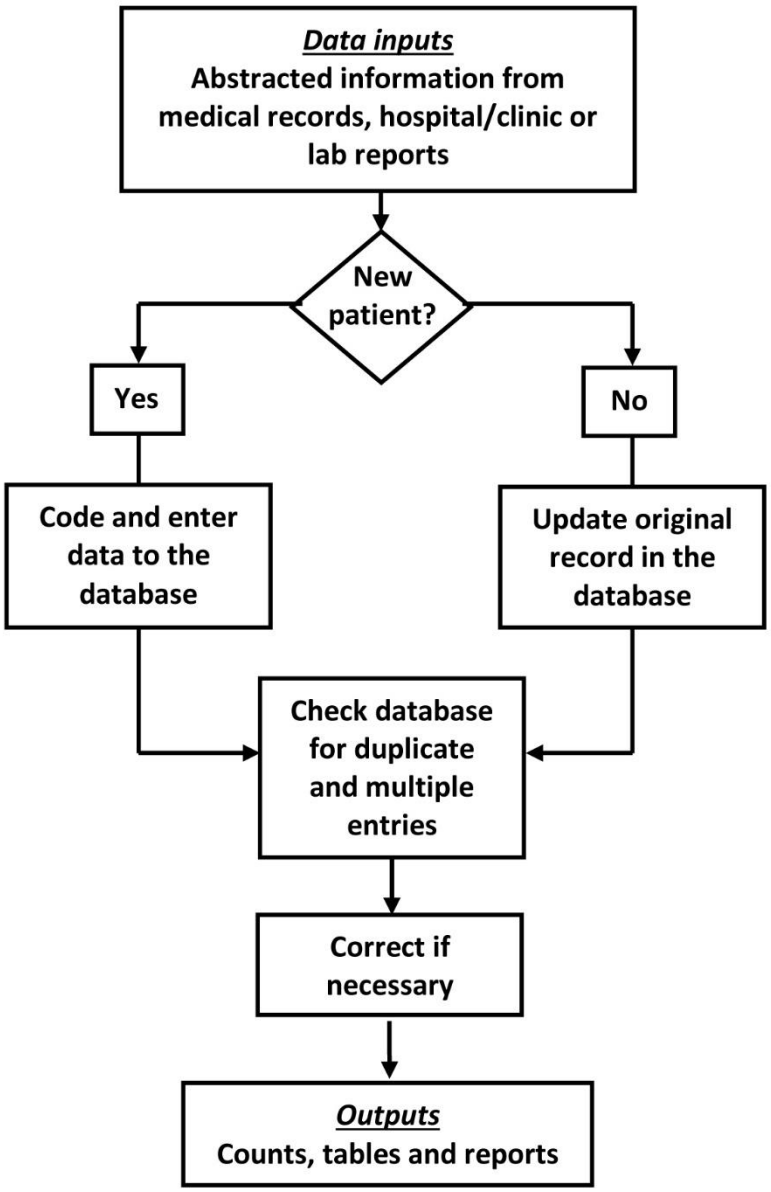


Fig 6. 3 Flow chart of data entry to database

DEATH CERTIFICATE NOTIFICATIONS

If the registration/notification form has been completed from a death certificate, leave **4 weeks** before data entry, to allow time to find the cases from any hospital sources.

Registration/notification forms completed from a death certificate (see section 4. ABSTRACTING) should be checked to see if the cancer case, or person, has already been registered, using the Browse / Edit function in CanReg in the same way as with a hospital abstract.

- ☞ If the case has been registered previously the record is updated with the date of death and any other new information. Death Certificate is added as a Souce of Information

- ☞ If there is no registration for the case; the place of death is checked. If the patient had died in hospital, the case should be "followed back" to see if the hospital record can be traced.

    - ➢ If it can be found - AND THE PATIENT REALLY DID HAVE CANCER - a registration/notification form should be completed from the hospital record with all the mandatory variables.
      <u>The case is registered with TWO sources (hospital and death certificate)</u>

    - ➢ If there has been no previous registration and it proves impossible to trace any record of the case having been seen in hospital:

EITHER

- The case is registered as a new cancer using the information on the death certificate.
- Enter basis of diagnosis = 0 (Death Certificate Only)
- Set date of incidence = date of death (UNLESS there is information on date of diagnosis on the certificate).
- Source of information will be death certificate.

OR

If there is doubt about the accuracy of the cause of death statement (for example, the certificate has been issued by a non-medical person), the case not registered (e.g. left pending).

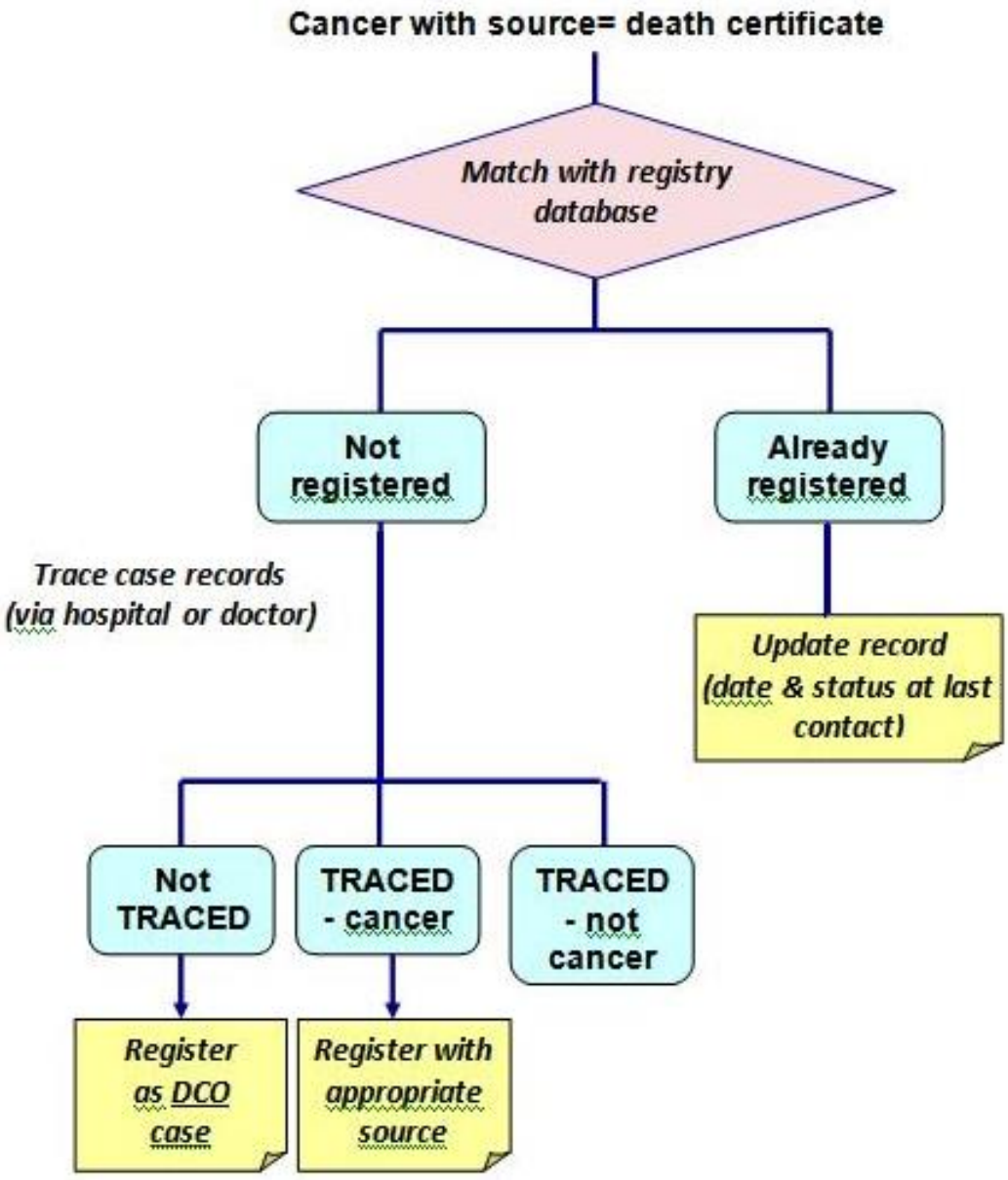


Fig 6. 4 The use of death certificates to identify new cases of cancer

7. DATA STORAGE

- Registration/notification forms must be filed numerically by registration number (Fig 7.1)
- Store in a cabinet that can be locked.
- The documents should be secure and inaccessible to unauthorized persons.
- They should be protected against loss or damage from fire, floods or any other interference.
- A BACK UP should be made of the CanReg database at the end of each day. The backup may be stored on a pen/flash drive/external portable hard drive/CD or other electronic media. This should be stored in a secure, locked cupboard or drawer that is secure and inaccessible to unauthorized persons.
- On transit data should be in a lockable suitcase.

Fig 7. 1 Filing of registration/notification forms

8. CONFIDENTIALITY

The Registry aims to maintain the confidentiality of all cancer information collected for the following reasons:

- To protect the privacy of the cancer patients
- To protect the privacy of the healthcare facilities reporting the cancer case
- To protect the privacy of the cancer patient's healthcare providers
- To protect from abuse and misuse of the cancer data

Guidelines on Confidentiality for Population-Based Cancer Registration have been published by IACR/IARC (IARC Internal Report No. 2004/03). It is available at http://www.iacr.com.fr/images/doc/confidentiality2004.pdf.

Definition of confidential data

Confidential data include data that identifies specific information on the patient, healthcare facilities and healthcare providers reporting the case. The cancer registry should maintain the same standards of confidentiality as applicable to the confidentiality of medical records and clinician-patient relationship.

All staff must sign confidentiality document to preserve the anonymity of the registry data and not to divulge any information even after employment ceases. An example is provided in Appendix 4.

8.1 LOGISTICAL ASPECT OF CONFIDENTIALITY

The following measures to ensure confidentiality should be implemented:

Data collection

For data collected on registration/notification forms, it is the responsibility of the registry staff to preserve their confidentiality. Forms should be kept under lock and key preferably in a filing cabinet. They should not be left in a place where an unauthorized person might have access e.g. in your car.

Data transmission

When sending information through the mail:

- Use registered mail.
- Information should be sent in two separate lists; one of names and the other medical information which when in the registry are merged.
- Use double envelopes; the exterior one with a general address and the interior with the address to the authorized recipient who should ideally be the registry director or delegated/authorised person.

Confidential data should **NEVER** be sent by fax.

When information is sent electronically such as USB hard drives or CDs it is important to take measures to ensure that these will not get lost, and not be easily read by other parties. The following precautions may be taken:

- Encrypting of the names at various level.
- Preparation of a separate CDs or USB with the names and one with the tumour related data.
- Keep a record of all electronically transmitted and received data.
- Data not to leave the registry premises without authorization.

Computer

With data kept in computer, usernames and passwords should be used and changed regularly and it should be known only to the authorised users.

Telephone

Confidential information should **NEVER** be given over the telephone, nor should enquiries from collaborators concerning confidential data be given over the telephone.

8.2 ACCESS TO AND STORAGE OF DATA

Strict security measures should be exercised to ensure confidentiality. These include:

- Access to the registry should be limited and restricted to authorized persons only (Fig 8.1).
- All registry records should be stored in a room which can be locked and access limited only to authorized persons.
- Provide lockable filing cabinets
- Use shredder machine to destroy unwanted forms

Fig 8. 1 Restricted Entry Notice

8.3 USE AND RELEASE OF DATA

Confidential data may be provided by the registry only upon written request, (see Appendix 5) which should include the exact purpose for which the data will be used, the information required the name(s) of the person(s) responsible for keeping the confidential information and the time period for which the data are needed.

- The registry should make sure that those receiving the data:
    - Are bound by the same rules of confidentiality observed by the registry staff.
    - Will use the data only for the purpose agreed upon at the time of provision, and will not make them accessible to other parties.
    - Will destroy the data when they are no longer needed for the said purposes.
- No information should be provided to insurance companies, medical funds pension schemes employers, the police or to a physician having to examine an individual for such purpose.

Aggregate data

These kinds of data do not need strict confidentiality measures and include prepared tables, graphs and reports.

Individual data

Cancer registries contribute to investigations on the cause of cancer and the registry may frequently be asked to provide the names of patients with given cancers so that they can be included in a study. Patients' names may be disclosed to the treating physician. Otherwise, patients' names can be disclosed to the researchers who have the authority/approval of the registry director and the ethics committee. Names may be disclosed to researchers with the agreement that the patients or members of the family may not be identified or any detailed information which permits any form of identification.

International release

When sending data abroad the registry staff should ensure that patients' identifications are not disclosed. Cases may be identified by a code number or patients' registry number (which can be linked to the registration record within the registry).

Requests by researchers for data from AFCRN members in more than one country should be referred to the AFCRN Research Committee.

8.4 DISSEMINATION OF DATA TO ORGANISATIONS OUTSIDE RESEARCH for example THE PRESS

Only the registry director may release of data to the media. He/she should insist on viewing the draft of the article prior to release or publication. Identifiable data should **NEVER** be released to the media.

9. QUALITY CONTROL

The primary goal of a population–based cancer registry is to determine the incidence of cancer within its geographical population. It is therefore of the utmost importance that the registry data be of good quality. This means that the information gathered, especially on essential items should be complete, consistent and accurate, and that coverage of the population should be as complete as possible. Quality control concerns three aspects of registry work:

VALIDITY: This is the accuracy of the information registered (or, the proportion of cases recorded as having a given characteristic that truly have that attribute).

COMPLETENESS: This is the extent to which all of the new (incident) cancers occurring in the target population of the registry are included in the database.

TIMELINESS: the speed with which registry data is ready for analysis and reporting.

9.1 MEASURING VALIDITY (ACCURACY) OF REGISTRY DATA

The methods used are as follows:

1. Re-abstracting and recoding "audits"
2. Reporting "Morphology Verified" percentages
3. Reporting DCO percentages
4. Reporting on percentage of missing information
5. Internal consistency checks

9.1.1 Re-abstracting and recoding audits

Re-abstracting audits and recoding audits often are used to assess the accuracy (agreement with source medical records) and reproducibility (agreement among data collectors) of registry data.

They need to be performed by an auditor – either from the registry (for example, the Director, or Registry Manager), or an "expert" consultant from outside.

The objective of a re-abstracting study is to measure the level of agreement between data in the registry and data re-abstracted and recoded by the auditor from source records (the hospital medical records for most cases).

Re-abstracting

A sample of registrations is selected from the registry database by the auditor. Eligible cases are those diagnosed at least one year prior to the year of the study.

He/she will select the sample:

- at random from the whole database
- randomly from certain sources that are known to cause problems to the registry staff
- randomly, but with the same number of cases drawn for each registrar

The sample will be for registrations from a single year (or period of a few years) that are subject of the quality control exercise.

Hilsenbeck et al, (1987), of the Centralized Cancer Patient Data System in the USA suggested that the sample size should be, as a minimum, 3-4 cases per registrar per month.

For these registrations, the records from which the case was abstracted are requested from the source concerned. This means sending a list of the case records required (the list contains case number, patient's name, date) to the sources (hospital records departments, for example) and requesting that the case files are ready for the exercise.

The auditor will then abstract the case onto the registration form (WITHOUT looking at the original registration). The re-abstracts are compared with the original (either the registration form, or the details from the CanReg database).

For each re-abstracted data item, the auditor's codes are compared to the original codes to identify discrepancies. If the codes do not match, the discrepancy is classified as to severity according to major and minor. Table 9.1 shows discrepancies that can be considered "Major" – any other discrepancy is considered "minor". The Table 9.2 shows an example of results of such a study.

Item	**Major disagreement**	**Item**	**Major disagreement**
Demographic		***Treatment***	
Sex	any difference	Type: surgery radiotherapy chemotherpy hormone therapy	given vs not given
Age	>1 years difference		
Birthdate	different year		
Ethnic group	any difference		
Place of residence	in/out of registry area	Date	difference ≥ 1 month
Tumour		***Follow up***	
Date of incidence	different year	Date of last contact	difference ≥ 3 months
Primary site	difference in first two digits of ICD-O C code	Status at last contact	any difference
Morphology	difference in first three digits of ICD-O M code		
Behaviour	any difference in ICD-O behaviour code		
Basis of diagnosis	difference MV or non-MV or DCO		
Stage	difference resulting in change of UICC stage (I-IV)		

Table 9. 1 Major disagreements for selected key data items – other discrepancies consided "minor"

Data Items	Data Items Reabstracted	Number in agreement	% agreement
Sex	50	50	100%
Ethnic goup	50	48	96%
Age	50	47	94%
Date of Diagnosis	50	43	86%
Primary Site	50	46	92%
Histology	50	46	92%
Basis of diagnosis	50	48	96%
Stage	50	33	66%
Treatment			
Surgery	50	48	96%
Radiation Therapy	50	47	94%
Chemo-Endocrine Therapy	50	46	92%
Other Therapy	50	50	100%
Date of Treatment	50	45	90%
Date of Last Contact	50	48	96%
Vital Status at Last Contact	50	49	98%
TOTALS	750	694	93%

Table 9. 2 Results of a Hypothetical Re-abstracting Study.

Recoding audits These look at the level of agreement between registry staff and the auditor for records already in the registry. The auditor uses the text contained on the registration form to recode a sample of actual case records in the registry database.

As in a re-abstracting study, for each recoded case, codes for each data item are compared for discrepancies with those assigned by the auditor. These studies show:

- The types of tumour records in which discrepancies occur more frequently.
- Sources of variation (e.g., misinterpretation of source document information, information not available at initial abstracting, misinterpretation of coding rules, inadequate or erroneous consolidation of data between records).
- Effect of misclassification that could affect data analysis and use (e.g., are tumours more frequently over-staged or under-staged?).
- Data quality with respect to other factors such as who collects the data (permanent registrars versus medical staff), training and skills of the registrars collecting the data, and difficulty of abstracting and coding the specific data items.

This information should be used to identify training needs and to modify registry processes and procedures to ensure future improvement in data quality.

9.1.2 Percentage of cases with a morphologically verified diagnosis (MV%)

Morphological verification refers to cases for which the diagnosis is based on histology or cytology.
The diagnostic information on cases that are morphologically verified is traditionally considered as a sort of "gold standard", with suspicion falling upon the accuracy of diagnosis by other means (although in reality a diagnosis based on an MRI or CT scan may be just as accurate as one based on exfoliative cytology). A high MV% is taken to mean accuracy of diagnosis, whereas a low MV% casts doubt on the validity of the data.

Procedure:

For the time period for which the quality control exercise is being performed (for example, one year, three years, 5 years), make a table, with, for each sex, the number of cases , by cancer site (using the ICD-10 codes) for each "Basis of Diagnosis" code (see Table 9.3, left side).

Then, group together the "basis of diagnosis codes" that represents diagnoses based on examination by microscope (generally in pathology or haematology labs). The codes (section 5.4, page 26) are:

5. Cytology or haematology
6. Histology of a metastasis
7. Histology of a primary tumour

The MV% is the percentage of all registrations with these "basis" codes.

The right hand side of Table 9.3 shows how the codes (ICD-10) for cancer site, and for "basis of diagnosis" can be grouped (with Basis of diagnosis as DCO/ Clinical/ M.V.) in a table suitable for publication in a registry report.

Basic data: Table of site (ICD-10) by basis of diagnosis

SITE	BASIS CODES										BASIS	
(ICD-10)	0	1	2	3	4	5	6	7	8	Total	1-4	5-8
0	0	3	0	0	0	0	0	0	0	3	3	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	4	0	0	0	0	0	11	0	15	4	11
3	0	1	0	1	0	0	0	5	0	7	2	5
4	0	0	0	0	0	0	0	0	0	0	0	0
5	0	4	0	0	0	0	0	6	0	10	4	6
6	1	5	0	1	0	0	0	3	0	10	6	3
7	0	5	0	0	0	0	0	6	0	11	5	6
8	0	1	0	0	0	0	0	3	0	4	1	3
9	0	2	0	0	0	0	0	5	0	7	2	5
10	0	4	0	0	0	0	0	6	0	10	4	6
11	0	22	0	0	0	0	1	27	0	50	22	28
12	0	0	0	0	0	0	0	1	0	1	0	1
13	0	2	0	0	0	0	0	1	0	3	2	1
14	0	2	0	0	0	0	0	0	0	2	2	0
15	16	95	0	0	0	0	0	71	0	182	95	71
16	8	33	0	1	0	0	2	37	0	81	34	39
17	0	0	0	0	0	0	0	3	0	3	0	3
18	1	27	0	0	0	0	0	22	0	50	27	22
19	0	7	0	0	0	0	0	5	0	12	7	5
20	4	24	0	0	0	0	0	34	0	62	24	34
21	0	2	0	0	0	0	0	4	0	6	2	4
22	8	104	1	0	0	0	0	87	0	200	105	87
23	0	0	0	0	0	0	0	4	0	4	0	4
25	3	27	0	0	0	0	0	6	0	36	27	6
26	0	0	0	0	0	0	0	2	0	2	0	2
30	0	8	0	0	0	0	0	11	0	19	8	11
31	0	1	0	0	0	0	0	2	0	3	1	2
32	0	13	0	0	0	0	0	10	0	23	13	10
34	2	37	0	0	0	0	0	20	0	59	37	20
37	0	3	0	0	0	0	0	0	0	3	3	0

Cancer site	ICD-10	Basis of diagnosis				
		No. Cases	(% total)	DCO	Clinical	M.V.
Oral cavity & pharynx	C00-C14	133	3.1	0.8%	42.9%	56.4%
Oesophagus	C15	182	4.3	8.8%	52.2%	39.0%
Stomach	C16	81	1.9	9.9%	42.0%	48.1%
Large bowel	C18-C21	130	3.1	3.8%	46.2%	50.0%
Liver	C22	200	4.7	4.0%	52.5%	43.5%
Pancreas	C25	36	0.9	8.3%	75.0%	16.7%
Larynx	C32	23	0.5	0.0%	56.5%	43.5%
Lung	C33-C34	59	1.4	3.4%	62.7%	33.9%
Bone	C40-C41	54	1.3	0.0%	42.6%	57.4%
Melanoma of Skin	C43	25	0.6	0.0%	28.0%	72.0%
Other Skin	C44	54	1.3	0.0%	37.0%	63.0%
Kaposi sarcoma	C46	1035	24.4	1.1%	25.2%	73.7%
Breast	C50	334	7.9	0.9%	53.3%	45.8%
Cervix Uteri	C53	492	11.6	2.6%	43.5%	53.9%
Corpus Uteri	C54	33	0.8	3.0%	27.3%	69.7%
Ovary	C56	57	1.3	5.3%	49.1%	45.6%
Prostate	C61	236	5.6	1.3%	33.9%	64.8%
Kidney	C64	55	1.3	0.0%	38.2%	61.8%
Bladder	C67	22	0.5	0.0%	50.0%	50.0%
Eye	C69	122	2.9	0.0%	15.6%	84.4%
Brain, Nervous system	C70-C72	36	0.9	8.3%	61.1%	30.6%
Thyroid	C73	35	0.8	2.9%	14.3%	82.9%
Hodgkin disease	C81	46	1.1	0.0%	15.2%	84.8%
Non-Hodgkin lymphom	C82-C85;C	298	7.0	1.3%	45.0%	53.7%
Leukaemia	C91-C95	98	2.3	6.1%	40.8%	53.1%
All sites Total	All	4235	100	2.2%	39.2%	58.6%

Table 9. 3 Example of calculation of MV% (Registry X, data for 2005-2007)

One of the standard tables in CanReg5 ("Data Quality Indicators") includes the MV% - in addition to other indicators of data quality (see Table 9.4).

Training System (English) (2001–2005)

Data Quality Indicators

MALE

SITE	Cases	% Total	ASR(se)	MV(%)	CLIN(%)	DCO(%)	ICD10
Mouth & pharynx	418	5.47	16.62 (0.84)	98.80	0.72	0.48	C00–14
Oesophagus	197	2.58	8.31 (0.61)	88.83	2.03	9.14	C15
Stomach	430	5.63	19.11 (0.95)	91.63	2.09	6.28	C16
Colon, rectum, anus	529	6.93	22.45 (1.01)	93.38	1.32	5.29	C18–21
Liver	116	1.52	5.09 (0.49)	60.34	8.62	31.03	C22
Pancreas	81	1.06	3.60 (0.41)	60.49	11.11	28.40	C25
Larynx	163	2.13	7.11 (0.57)	95.09	3.68	1.23	C32
Lung, trachea, bronchus	500	6.55	23.14 (1.05)	79.80	6.60	13.60	C33–34
Pleura & other thoracic	20	0.26	0.67 (0.16)	70.00	15.00	15.00	C37–38
Melanoma of skin	122	1.60	4.84 (0.46)	96.72	0.82	2.46	C43
Prostate	2153	28.19	105.44 (2.29)	95.45	1.49	3.07	C61
Testis	56	0.73	1.41 (0.20)	92.86	3.57	3.57	C62
Kidney & urinary NOS	132	1.73	5.52 (0.50)	91.67	2.27	6.06	C64–66,68
Bladder	265	3.47	12.42 (0.77)	94.34	2.26	3.40	C67
Brain & nervous sytem	211	2.76	7.16 (0.53)	82.46	5.21	12.32	C70–72
Thyroid	65	0.85	2.08 (0.28)	98.46	0.00	1.54	C73
Ill–defined	204	2.67	8.92 (0.64)	63.24	12.25	24.51	C76–80
Lymphoma	414	5.42	15.93 (0.82)	87.44	2.17	10.39	C81–85,90,88,96
Leukaemia	226	2.96	7.91 (0.56)	76.55	1.33	22.12	C91–95
All sites but C44	6727	88.08	293.72 (3.70)	89.58	2.76	7.66	ALLbC44

FEMALE

SITE	Cases	% Total	ASR(se)	MV(%)	CLIN(%)	DCO(%)	ICD10
Mouth & pharynx	139	1.79	4.45 (0.39)	94.24	3.60	2.16	C00–14
Oesophagus	61	0.78	2.20 (0.29)	90.16	1.64	8.20	C15
Stomach	278	3.58	9.40 (0.58)	92.81	2.16	5.04	C16
Colon, rectum, anus	608	7.82	20.70 (0.86)	93.91	1.48	4.61	C18–21
Liver	53	0.68	1.80 (0.25)	47.17	3.77	49.06	C22
Pancreas	104	1.34	3.68 (0.37)	57.69	9.62	32.69	C25
Larynx	36	0.46	1.28 (0.22)	97.22	2.78	0.00	C32
Lung, trachea, bronchus	308	3.96	11.11 (0.64)	78.25	5.84	15.91	C33–34
Pleura & other thoracic	11	0.14	0.31 (0.10)	81.82	0.00	18.18	C37–38
Melanoma of skin	124	1.60	3.80 (0.36)	100.00	0.00	0.00	C43
Breast	1766	22.72	54.78 (1.35)	97.06	1.30	1.64	C50
Cervix	904	11.63	26.70 (0.93)	98.23	0.55	1.22	C53
Corpus & Uterus NOS	226	2.91	7.78 (0.53)	98.23	0.88	0.88	C54–55
Ovary & adnexa	228	2.93	7.24 (0.50)	92.11	2.19	5.70	C56
Kidney & urinary NOS	93	1.20	3.30 (0.35)	92.47	2.15	5.38	C64–66,68
Bladder	121	1.56	4.29 (0.40)	91.74	1.65	6.61	C67
Brain & nervous sytem	159	2.05	4.61 (0.38)	78.62	4.40	16.98	C70–72
Thyroid	366	4.71	9.99 (0.55)	96.99	1.64	1.37	C73
Ill–defined	173	2.23	5.88 (0.46)	70.52	10.40	19.08	C76–80
Lymphoma	380	4.89	11.98 (0.64)	89.21	2.11	8.68	C81–85,90,88,96
Leukaemia	234	3.01	6.93 (0.47)	84.62	0.43	14.96	C91–95
All sites but C44	6750	86.85	214.37 (2.70)	92.00	2.09	5.91	ALLbC44

Cases of unknown age (21 M / 26 F) were excluded from these analyses

Table 9. 4 Output of CanReg-5 (Data Quality Indicators)

The absolute value of the MV% needs to be compared with an “expected” value that is reasonable given the circumstances (state of medical technology, local clinical practice) in which the registry operates. Therefore, the

MV values (by site and, preferably also by sex) should be compared with an appropriate set of standards, so that values that are significantly different can be identified.

Table 9.5 provides the "standard" values of MV% for sub-Saharan Africa, with which your own values can be compared[5].

ICD-10 code	Cancer site	Male	Female
		MV%	MV%
C00–14	Oral cavity and pharynx	68.6	71.4
C15	Oesophagus	46.7	45.9
C16	Stomach	53.1	53.4
C18–21	Large bowel	62.1	61.3
C22	Liver	11.7	12.6
C25	Pancreas	16.8	22.2
C32	Larynx	66.2	73.3
C33–34	Trachea, bronchus, and lung	44.7	64.1
C43	Melanoma of skin	76.9	90.0
C50	Breast	66.7	66.1
C53	Cervix uteri	0.0	62.4
C54–55	Corpus uteri, uterus unspecified	0.0	64.6
C56	Ovary	0.0	51.3
C61	Prostate	59.8	0.0
C62	Testis	48.3	0.0
C64–66	Kidney, renal pelvis, and ureter	68.8	67.1
C67	Bladder	39.7	45.0
C70–72	Brain, central nervous system	51.5	41.8
C73	Thyroid	65.4	73.8
C81–88, C90	Lymphomas	84.5	82.0
C91–95	Leukaemia	87.2	88.4
C76–80	Unspecified	48.4	39.8
C00–96 (excluding C44)	All sites (excluding non-melanoma skin)	57.4	61.1

MV%, percentage of cases with a morphologically verified diagnosis.
[a] The Gambia (1997–1998), Mali, Bamako (1994–1996), Uganda, Kyadondo County (1993–1997), Zimbabwe, Harare: African (1993–1997).

Table 9. 5 Mean values of MV% for cancer registries in sub-Saharan Africa (from IARC Technical Report 43 (2014))

The CanReg5 Table ("Data Quality Indicators") does not yet show whether the recorded MV% is significantly different from (higher or lower) this standard (see Table 9.4).

Whereas a MV% significantly lower than the expected value may give rise to concern about a lack of validity, it is generally not the cancer registry that can influence the availability of, or use of, pathology services within its area. Usually, in Africa, the opposite situation – a relatively high MV% – is cause for concern. This is because collecting data on cancer cases from pathology departments is much easier than trawling through clinical services or ill-organized hospital archives. A large proportion of cases diagnosed via the pathology department may well suggest defects in case finding and, hence, incomplete registration. Worse, the incompleteness will be biased, with the

[5] A suitable statistical test is has been described in Bray & Parkin (2009)

database containing a deficit of cancers that are not easy to biopsy, and so are diagnosed by other methods (e.g. lung, liver, brain, and pancreatic cancer).

9.1.3 Percentage of cases for which the only information came from a death certificate (DCO%)

DCO cases are those registered on the basis of information on a death certificate, and for which no other information could be traced. As described earlier (section 6.1), the nature of death certificates in Africa varies widely, from those issued as part of a civil registration of vital events to those generated in a hospital mortuary.

However, almost always the accuracy of the diagnostic information is questionable, since the person writing out the certificate may have had little contact with the patient before death and may be ill-informed about how to record cause of death. They may even have no medical training at all. Thus, if no other clinical record for persons who apparently died of (or with) cancer can be found, there is a reasonable suspicion that the diagnosis was simply wrong.

If you include such cases in the database, and if they comprise a large proportion of cases, the validity of the data is suspect.

Procedure:

As for MV% (see 9.1.2), for the time period for which the quality control exercise is being performed (for example, one year, three years, 5 years), make a table, with, for each sex, the number of cases, by cancer site (using the ICD-10 codes) for each "Basis of Diagnosis" code.

The DCO cases are those with *basis of diagnosis* = 0
See Table 9.3

The DCO% is the percentage of all registrations with this "basis" code (=0)
As for MV%, we calculate DCO% by cancer site, and, ideally, by sex.

The CanReg5 Table (Data Quality Indicators) shows the percentage of DCO cases, by site and sex (see Table 9.4).

What is an Acceptable Level of DCO% ?

This is difficult – it depends on local circumstances, for example availability of death certificates, success in record linkage, accuracy of cause of death statements on the certificate.

Some collections of cancer registry results have proposed more or less arbitrary standards; for example, Cancer Incidence in Five Continents Volume IX (Curado et al, 2007) considered <10% DCO to be category "A" for quality, and 10-20% category "B". The criteria in the North American Association of Central Cancer Registries (NAACCR) is a DCO of less than 3 percent for "gold" standard, and less than 5 percent for silver (Hofferkamp, 2008).

9.1.4 Proportion (or percentage) of cases with missing data

The proportion of cases with unknown values of different data items, such as age or stage, is also an indicator of data quality. The data items that should be assessed for missing values are:

- Age
- Primary site
- Stage
- Follow up

Note that it is **NEVER** acceptable for "sex" to be missing.

Unknown values can result from problems with the registration process, may also result from inadequate case histories or investigation, or ambiguity in the medical record.

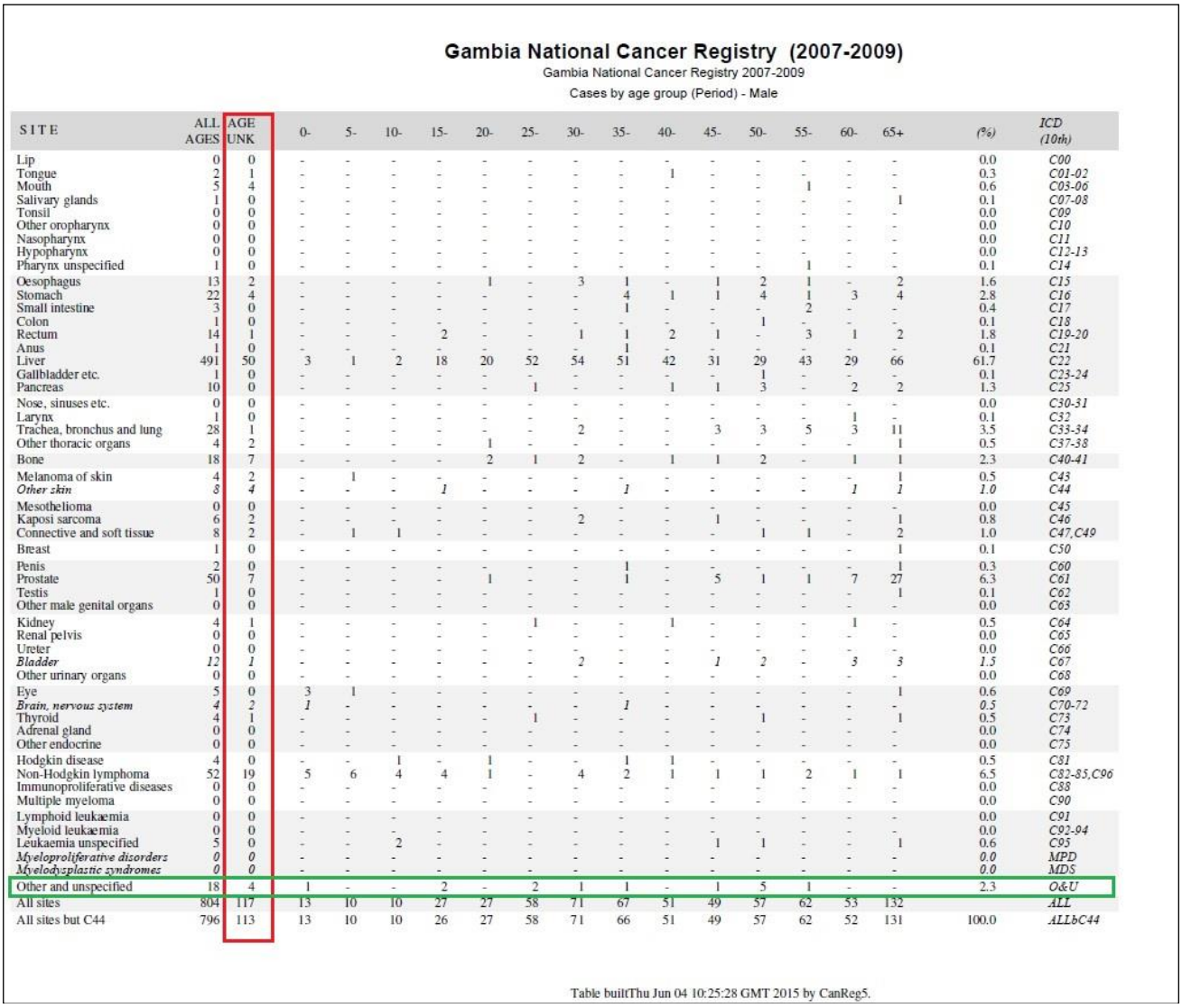
For "AGE" we wish to calculate the number of registrations (by cancer site and sex) for which age was recorded as unknown (code 99).

Primary site uncertain (PSU%) includes, in addition to "Unknown Primary Site" (C80 in ICD-10), other rubrics (e.g. malignant neoplasms of ill-defined organs of the digestive system (ICD-10 C26), respiratory system (C39) endocrine system (C75), and peritoneal and retroperitoneal neoplasms (C48) as well as those of "Other and Ill-defined Sites" (C76).

The standard Tables produced by CanReg show the numbers of cases with Age Unknown, and with a row entitled "Other and Unspecified" (O & U) (Fig 9.6).

A high proportion of cases assigned to the O&U/PSU category means there is low accuracy of diagnosis, usually due to the failure to specify the site of the primary cancer in cases diagnosed on the basis of tissue obtained from a metastasis. Incidence rates for cancers at specific sites will be underestimated if a large proportion of registered cases appear in the "Other and/or Unspecified" category, rather than with their true diagnosis.

As for DCO%, some collections of cancer registry results have proposed more or less arbitrary standards for % missing; for example, CI5 Volume IX proposed acceptable maxima for the percentage of cases with age unknown (<20%), and ill defined sites (<20%) (Curado et al, 2007). The NAACCR standards include <3% cases with Age missing and <5% cases with unknown primary site (Hofferkamp, 2008).



Gambia National Cancer Registry (2007-2009)

Gambia National Cancer Registry 2007-2009

Cases by age group (Period) - Male

SITE	ALL AGES	AGE UNK	0-	5-	10-	15-	20-	25-	30-	35-	40-	45-	50-	55-	60-	65+	(%)	ICD (10th)
Lip	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C00
Tongue	2	1	-	-	-	-	-	-	-	-	1	-	-	-	-	-	0.3	C01-02
Mouth	5	4	-	-	-	-	-	-	-	-	-	-	-	1	-	-	0.6	C03-06
Salivary glands	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.1	C07-08
Tonsil	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C09
Other oropharynx	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C10
Nasopharynx	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C11
Hypopharynx	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C12-13
Pharynx unspecified	1	0	-	-	-	-	-	-	-	-	-	-	-	1	-	-	0.1	C14
Oesophagus	13	2	-	-	-	-	1	-	3	1	-	1	2	1	-	2	1.6	C15
Stomach	22	4	-	-	-	-	-	-	-	4	1	1	4	1	3	4	2.8	C16
Small intestine	3	0	-	-	-	-	-	-	-	1	-	-	-	2	-	-	0.4	C17
Colon	1	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	0.1	C18
Rectum	14	1	-	-	-	2	-	-	1	1	2	1	-	3	1	2	1.8	C19-20
Anus	1	0	-	-	-	-	-	-	-	1	-	-	-	-	-	-	0.1	C21
Liver	491	50	3	1	2	18	20	52	54	51	42	31	29	43	29	66	61.7	C22
Gallbladder etc.	1	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	0.1	C23-24
Pancreas	10	0	-	-	-	-	-	1	-	-	1	1	3	-	2	2	1.3	C25
Nose, sinuses etc.	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C30-31
Larynx	1	0	-	-	-	-	-	-	-	-	-	-	-	-	1	-	0.1	C32
Trachea, bronchus and lung	28	1	-	-	-	-	-	-	2	-	-	3	3	5	3	11	3.5	C33-34
Other thoracic organs	4	2	-	-	-	-	1	-	-	-	-	-	-	-	-	1	0.5	C37-38
Bone	18	7	-	-	-	-	2	1	2	-	1	1	2	-	1	1	2.3	C40-41
Melanoma of skin	4	2	-	1	-	-	-	-	-	-	-	-	-	-	-	1	0.5	C43
Other skin	*8*	*4*	-	-	-	*1*	-	-	-	*1*	-	-	-	-	*1*	*1*	*1.0*	C44
Mesothelioma	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C45
Kaposi sarcoma	6	2	-	-	-	-	-	-	2	-	-	1	-	-	-	1	0.8	C46
Connective and soft tissue	8	2	-	1	1	-	-	-	-	-	-	-	1	1	-	2	1.0	C47,C49
Breast	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.1	C50
Penis	2	0	-	-	-	-	-	-	-	1	-	-	-	-	-	1	0.3	C60
Prostate	50	7	-	-	-	-	1	-	-	1	-	5	1	1	7	27	6.3	C61
Testis	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.1	C62
Other male genital organs	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C63
Kidney	4	1	-	-	-	-	-	1	-	-	1	-	-	-	1	-	0.5	C64
Renal pelvis	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C65
Ureter	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C66
Bladder	*12*	*1*	-	-	-	-	-	-	*2*	-	-	*1*	*2*	-	*3*	*3*	*1.5*	C67
Other urinary organs	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C68
Eye	5	0	3	1	-	-	-	-	-	-	-	-	-	-	-	1	0.6	C69
Brain, nervous system	*4*	*2*	*1*	-	-	-	-	-	-	*1*	-	-	-	-	-	-	*0.5*	C70-72
Thyroid	4	1	-	-	-	-	-	1	-	-	-	-	1	-	-	1	0.5	C73
Adrenal gland	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C74
Other endocrine	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C75
Hodgkin disease	4	0	-	-	1	-	1	-	-	1	1	-	-	-	-	-	0.5	C81
Non-Hodgkin lymphoma	52	19	5	6	4	4	1	-	4	2	1	1	1	2	1	1	6.5	C82-85,C96
Immunoproliferative diseases	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C88
Multiple myeloma	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C90
Lymphoid leukaemia	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C91
Myeloid leukaemia	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0	C92-94
Leukaemia unspecified	5	0	-	-	2	-	-	-	-	-	-	1	1	-	-	1	0.6	C95
Myeloproliferative disorders	*0*	*0*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*0.0*	MPD
Myelodysplastic syndromes	*0*	*0*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*0.0*	MDS
Other and unspecified	18	4	1	-	-	2	-	2	1	1	-	1	5	1	-	-	2.3	O&U
All sites	804	117	13	10	10	27	27	58	71	67	51	49	57	62	53	132		ALL
All sites but C44	796	113	13	10	10	26	27	58	71	66	51	49	57	62	52	131	100.0	ALLbC44

Table builtThu Jun 04 10:25:28 GMT 2015 by CanReg5.

Table 9. 6 Standard Table (Cases by age group (period)) from CanReg5 showing the percentage of cases with age unknown (AGE UNK) by site, and % with "Other and Unspecified" sites (by age group)

9.1.5. Consistency checks

In computerized registries some aspects of validity of registered data are checked using automated routines. This is done when the data are being entered into CANREG, or as a part of a batch operation (off-line). A 'scale of errors' is set up in the system such that major errors result in complete rejection of a registration, while less serious ones are flagged to indicate that they contain an error. These cases must be "*Confirmed*" in CanReg, otherwise, they remain as "*Pending*" cases, and will not appear in analytic tables.

The most basic edit check is on the validity of the codes used, so that records with coded values outside the permitted range for the item (as defined in the registry) are rejected.

At the next level are checks of logical consistency between data items. A cancer cannot be diagnosed before the date of birth of a patient, a man cannot have ovarian cancer, and treatment cannot be undertaken for a patient who has died.

An edit program rejects these impossible combinations. It may also flag unlikely or unusual combinations such as those site-specific morphology terms which have only one possible topography code e.g. nephroblastoma which arises from the kidney should have a topography code C64.9 and hepatoma which arises from the liver should have a topography code of C22.0.

The IARC-CHECK program is incorporated into CanReg, and also available to check files of cases "in batch" – that is, outside of CanReg - is available as part of the IARCcrgTools package:http://www.iacr.com.fr/index.php?option=com_content&view=category&layout=blog&id=68&Itemid=445
It checks data for validity of the following data items:

- registration number
- date of incidence
- age (or date of birth)
- sex
- site
- histology
- basis of diagnosis

and performs *data combination edits*, including:

- Incidence/birth dates
- Age/incidence/birth dates
- Age/site/histology
- Site/histology
- Sex/site
- Sex/histology
- Behaviour/site
- Behaviour/histology
- Basis of diagnosis/histology

The precise checks carried out are described in the manual "CHECK AND CONVERSION PROGRAMS FOR CANCER REGISTRIES (IARC/IACR Tools for Cancer Registries) J. Ferlay, C. Burkhard, S. Whelan, D.M. Parkin IARC Technical Report No. 42 Lyon, 2005, available at : http://www.iacr.com.fr/images/doc/TechRep42.pdf

If any of these checks on the data fail, CanReg gives warnings at the time of data entry. When the IARC-CHECK program is run in batch mode, it produces:

A. An output data file, which has the same layout as the input file but with symbol(s) and the new codes written at the end or replacing the original codes.
B. A warning file, created in the same directory as the input file, with the same name, but with the extension .CHK. It contains records which have been written to the output file, but which should be checked.
C. An error file, created in the same directory as the input file, with the same name the extension .ERR. The file contains all invalid combinations of items; these records are NOT included in the output file.

9.2 MEASURING COMPLETENESS OF REGISTRY DATA

The population-based registry aims to record all cancer cases occurring within its defined geographical area. It is therefore essential that all the data sources for the registry be covered completely. That is, case-finding and abstracting should include all hospitals within the catchment area of the registry. All data sources within these hospitals should likewise be covered in order to avoid under-reporting.

There are a number of methods that provide some indication of the completeness of a registry, but which do not actually quantify the number of cases missing. They include the following, discussed in more detail below:

9.2.1. Historic data methods
- 9.2.1.1 Stability of incidence rates over time
- 9.2.1.2 Comparison of incidence rates in different populations
- 9.2.1.3 Shape of age-specific curves
- 9.2.1.4 Incidence rates of childhood cancers

9.2.2. Mortality:Incidence ratios

9.2.3. Number of sources/notifications per case

Three methods are available to obtain a quantitative evaluation of the degree of completeness of registration:

9.2.4 Independent case ascertainment

9.2.5 Capture-recapture methods

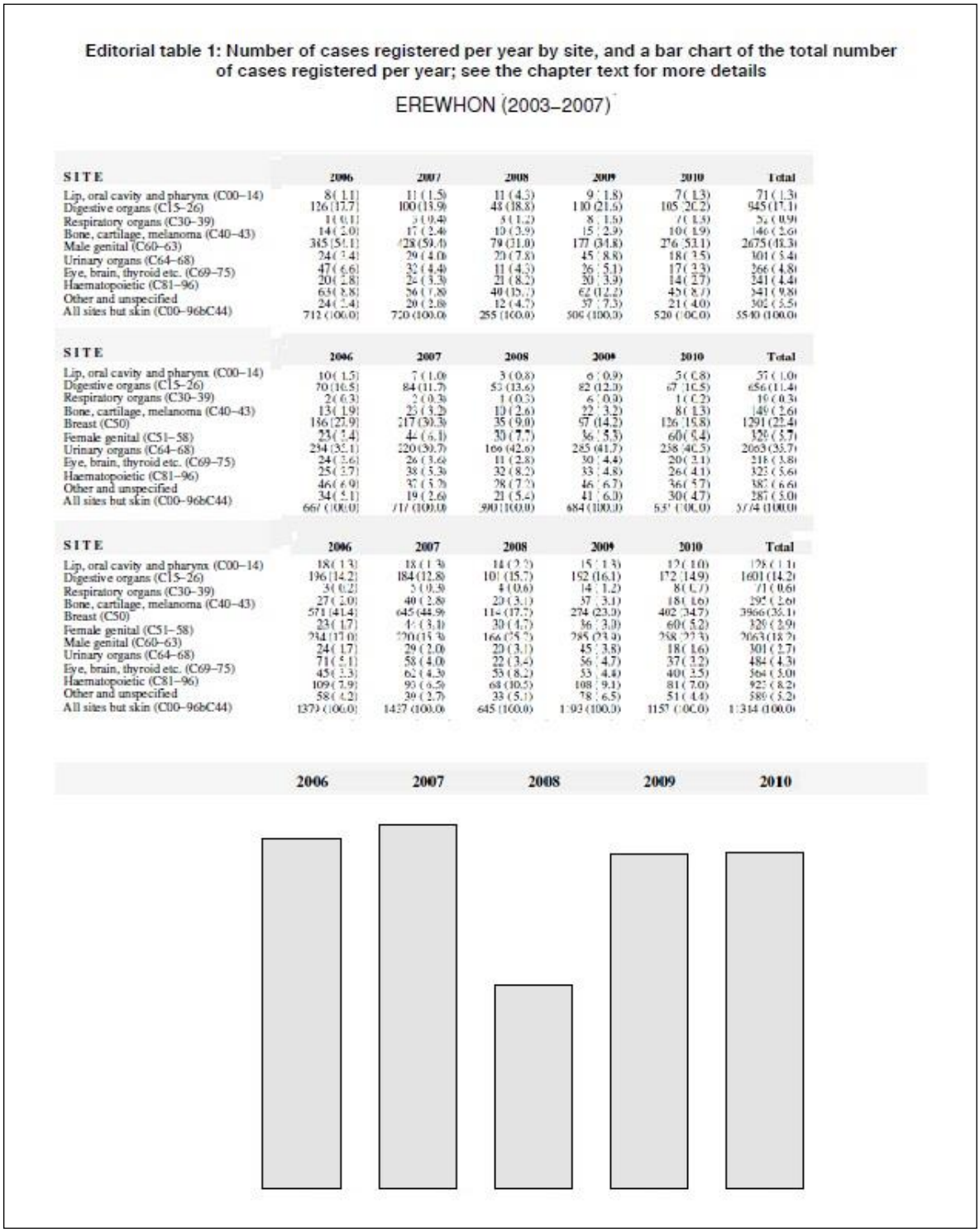
9.2.6 Death certificate methods

9.2.1. Historic data methods

9.2.1.1 Stability of incidence rates over time

If the registration area remains constant, then the number of cases registered per year might be expected to show only small and gradual changes from one year to the next. Quite often, the numbers of registrations will be increasing over time (as the population covered gets bigger, and older), so that looking at rates of incidence may be more useful.

However, it is very useful to use CanReg to prepare tables and graphs of time trends, by cancer type, and year, using the “Number of cases in major diagnosis groups in single calendar year of observation” option (Table 9.7).



Editorial table 1: Number of cases registered per year by site, and a bar chart of the total number of cases registered per year; see the chapter text for more details

EREWHON (2003–2007)

SITE	2006	2007	2008	2009	2010	Total
Lip, oral cavity and pharynx (C00–14)	8 (1.1)	11 (1.5)	11 (4.3)	9 (1.8)	7 (1.3)	71 (1.3)
Digestive organs (C15–26)	126 (17.7)	100 (13.9)	48 (18.8)	110 (21.5)	105 (20.2)	945 (17.1)
Respiratory organs (C30–39)	1 (0.1)	3 (0.4)	3 (1.2)	8 (1.5)	7 (1.3)	52 (0.9)
Bone, cartilage, melanoma (C40–43)	14 (2.0)	17 (2.4)	10 (3.9)	15 (2.9)	10 (1.9)	146 (2.6)
Male genital (C60–63)	365 (51.1)	428 (59.4)	79 (31.0)	177 (34.8)	276 (53.1)	2675 (48.3)
Urinary organs (C64–68)	24 (3.4)	29 (4.0)	20 (7.8)	45 (8.8)	18 (3.5)	301 (5.4)
Eye, brain, thyroid etc. (C69–75)	47 (6.6)	32 (4.4)	11 (4.3)	26 (5.1)	17 (3.3)	266 (4.8)
Haematopoietic (C81–96)	20 (2.8)	24 (3.3)	21 (8.2)	20 (3.9)	14 (2.7)	241 (4.4)
Other and unspecified	63 (8.8)	26 (3.6)	40 (15.7)	62 (12.2)	45 (8.7)	541 (9.8)
All sites but skin (C00–96bC44)	24 (3.4)	20 (2.8)	12 (4.7)	37 (7.3)	21 (4.0)	302 (5.5)
	712 (100.0)	720 (100.0)	255 (100.0)	506 (100.0)	520 (100.0)	5540 (100.0)

SITE	2006	2007	2008	2009	2010	Total
Lip, oral cavity and pharynx (C00–14)	10 (1.5)	7 (1.0)	3 (0.8)	6 (0.9)	5 (0.8)	37 (1.0)
Digestive organs (C15–26)	70 (10.5)	84 (11.7)	53 (13.6)	82 (12.0)	67 (10.5)	656 (11.4)
Respiratory organs (C30–39)	2 (0.3)	2 (0.3)	1 (0.3)	6 (0.9)	1 (0.2)	19 (0.3)
Bone, cartilage, melanoma (C40–43)	13 (1.9)	23 (3.2)	10 (2.6)	22 (3.2)	8 (1.3)	148 (2.6)
Breast (C50)	186 (27.9)	217 (30.3)	35 (9.0)	97 (14.2)	126 (19.8)	1291 (22.4)
Female genital (C51–58)	23 (3.4)	44 (6.1)	30 (7.7)	36 (5.3)	60 (9.4)	329 (5.7)
Urinary organs (C64–68)	234 (35.1)	220 (30.7)	166 (42.6)	285 (41.7)	238 (40.5)	2063 (35.7)
Eye, brain, thyroid etc. (C69–75)	24 (3.6)	26 (3.6)	11 (2.8)	30 (4.4)	20 (3.1)	218 (3.8)
Haematopoietic (C81–96)	25 (3.7)	38 (5.3)	32 (8.2)	33 (4.8)	26 (4.1)	323 (5.6)
Other and unspecified	46 (6.9)	37 (5.2)	28 (7.2)	46 (6.7)	36 (5.7)	382 (6.6)
All sites but skin (C00–96bC44)	34 (5.1)	19 (2.6)	21 (5.4)	41 (6.0)	30 (4.7)	287 (5.0)
	667 (100.0)	717 (100.0)	390 (100.0)	684 (100.0)	637 (100.0)	5774 (100.0)

SITE	2006	2007	2008	2009	2010	Total
Lip, oral cavity and pharynx (C00–14)	18 (1.3)	18 (1.3)	14 (2.2)	15 (1.3)	12 (1.0)	128 (1.1)
Digestive organs (C15–26)	196 (14.2)	184 (12.8)	101 (15.7)	192 (16.1)	172 (14.9)	1601 (14.2)
Respiratory organs (C30–39)	3 (0.2)	5 (0.3)	4 (0.6)	14 (1.2)	8 (0.7)	71 (0.6)
Bone, cartilage, melanoma (C40–43)	27 (2.0)	40 (2.8)	20 (3.1)	37 (3.1)	18 (1.6)	295 (2.6)
Breast (C50)	571 (41.4)	645 (44.9)	114 (17.7)	274 (23.0)	402 (34.7)	3966 (35.1)
Female genital (C51–58)	23 (1.7)	48 (3.3)	30 (4.7)	36 (3.0)	60 (5.2)	329 (2.9)
Male genital (C60–63)	234 (17.0)	220 (15.3)	166 (25.7)	285 (23.9)	258 (22.3)	2063 (18.2)
Urinary organs (C64–68)	24 (1.7)	29 (2.0)	20 (3.1)	45 (3.8)	18 (1.6)	301 (2.7)
Eye, brain, thyroid etc. (C69–75)	71 (5.1)	58 (4.0)	22 (3.4)	56 (4.7)	37 (3.2)	484 (4.3)
Haematopoietic (C81–96)	45 (3.3)	62 (4.3)	53 (8.2)	53 (4.4)	40 (3.5)	564 (5.0)
Other and unspecified	109 (7.9)	93 (6.5)	68 (10.5)	108 (9.1)	81 (7.0)	922 (8.2)
All sites but skin (C00–96bC44)	58 (4.2)	39 (2.7)	33 (5.1)	78 (6.5)	51 (4.4)	589 (5.2)
	1371 (100.0)	1427 (100.0)	645 (100.0)	1193 (100.0)	1157 (100.0)	11314 (100.0)

Table 9. 7 Number of cases registered per year by site, and a bar chart of the total number of cases registered per year (CanReg Table "Number of cases in major diagnosis groups in single calendar year of observation")

Irregular numbers of cases suggest that case-finding was imperfect for the periods concerned. In the example shown, there is a fall off in the numbers of cases registered in 2008, and, although this affects many cancer types, the numbers of Kaposi sarcoma cases, in particular, shows a dramatic fall in that year. Under-reporting may be site specific e.g. researchers may have carried out some study on a particular cancer and medical records on patients involved may be taken out by the researcher and the registry staff may not locate them.

CanReg5 also allows time trends of the incidence rates (age standardised) of the major cancers to be plotted as a graph (Fig 9.1).

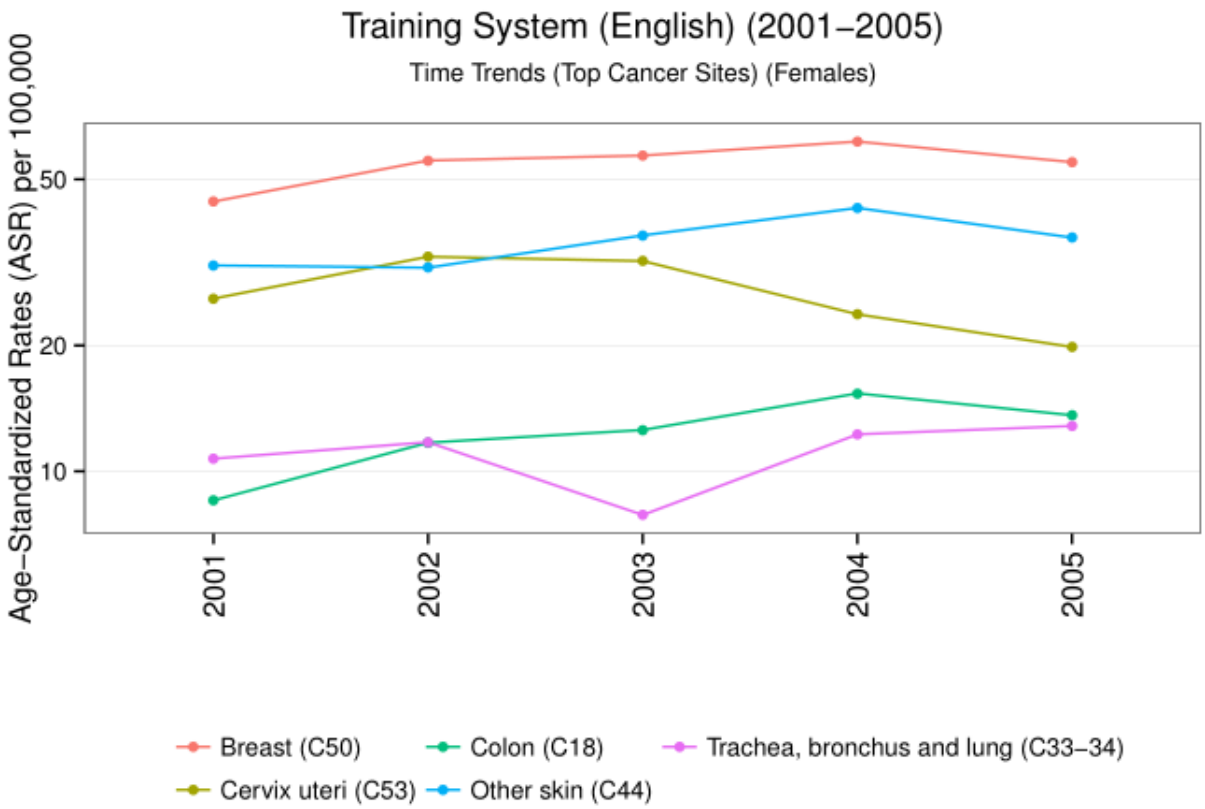


Fig 9. 1 Standard CanReg-5 output "Time trends (Top Cancer Sites)"

Another important check of registrations over time is to analyse the number of reports from different sources, by year of reporting.

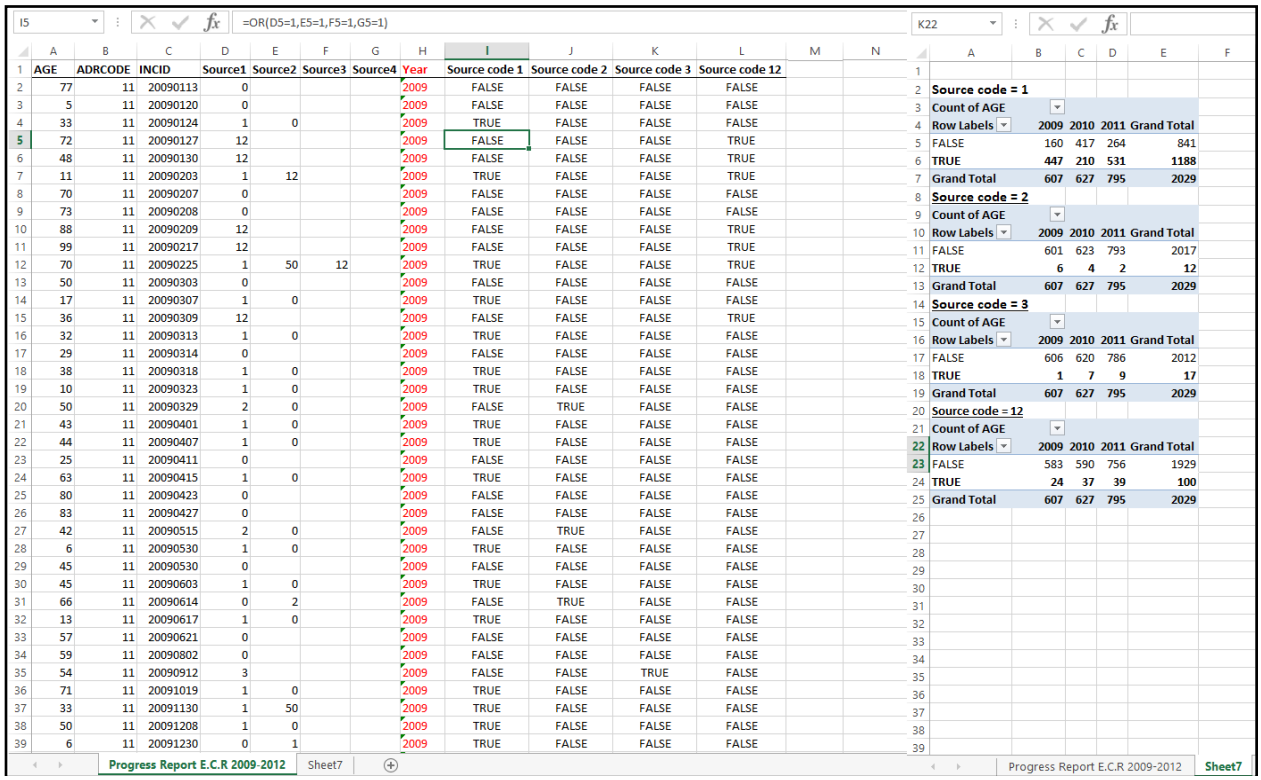
In CanReg, a single tumour can be recorded as being found in several different sources. However, at the time of publication, there was no standard analysis within CanReg to calculate the number of notifications of a single case from different sources.

Procedure:
Make an export file of cases for the group of interest – defined by the years (for which the cases were registered) and the place of residence of the cases (geographic area).

For each case, make sure that the date of incidence and the codes for all of the different SOURCES of information are present.

Table 9.8 (left side) shows an example of an export file. For each source code (for example, hospital, or laboratory), see if it is present in Source 1, or Source 2, or Source 3, or Source 4 etc for all of the cases in the file.

Calculate the numbers of notifications from that source in one year. Table 9.8 (right side) shows an example of the calculation, using EXCEL.



AGE	ADRCODE	INCID	Source1	Source2	Source3	Source4	Year	Source code 1	Source code 2	Source code 3	Source code 12
77	11	20090113	0				2009	FALSE	FALSE	FALSE	FALSE
5	11	20090120	0				2009	FALSE	FALSE	FALSE	FALSE
33	11	20090124	1	0			2009	TRUE	FALSE	FALSE	FALSE
72	11	20090127	12				2009	FALSE	FALSE	FALSE	TRUE
48	11	20090130	12				2009	FALSE	FALSE	FALSE	TRUE
11	11	20090203	1	12			2009	TRUE	FALSE	FALSE	TRUE
70	11	20090207	0				2009	FALSE	FALSE	FALSE	FALSE
73	11	20090208	0				2009	FALSE	FALSE	FALSE	FALSE
88	11	20090209	12				2009	FALSE	FALSE	FALSE	TRUE
99	11	20090217	12				2009	FALSE	FALSE	FALSE	TRUE
70	11	20090225	1	50	12		2009	TRUE	FALSE	FALSE	TRUE
50	11	20090303	0				2009	FALSE	FALSE	FALSE	FALSE
17	11	20090307	1	0			2009	TRUE	FALSE	FALSE	FALSE
36	11	20090309	12				2009	FALSE	FALSE	FALSE	TRUE
32	11	20090313	1	0			2009	TRUE	FALSE	FALSE	FALSE
29	11	20090314	0				2009	FALSE	FALSE	FALSE	FALSE
38	11	20090318	1	0			2009	TRUE	FALSE	FALSE	FALSE
10	11	20090323	1	0			2009	TRUE	FALSE	FALSE	FALSE
50	11	20090329	2	0			2009	FALSE	TRUE	FALSE	FALSE
43	11	20090401	1	0			2009	TRUE	FALSE	FALSE	FALSE
44	11	20090407	1	0			2009	TRUE	FALSE	FALSE	FALSE
25	11	20090411	0				2009	FALSE	FALSE	FALSE	FALSE
63	11	20090415	1	0			2009	TRUE	FALSE	FALSE	FALSE
80	11	20090423	0				2009	FALSE	FALSE	FALSE	FALSE
83	11	20090427	0				2009	FALSE	FALSE	FALSE	FALSE
42	11	20090515	2	0			2009	FALSE	TRUE	FALSE	FALSE
6	11	20090530	1	0			2009	TRUE	FALSE	FALSE	FALSE
45	11	20090530	0				2009	FALSE	FALSE	FALSE	FALSE
45	11	20090603	1	0			2009	TRUE	FALSE	FALSE	FALSE
66	11	20090614	0	2			2009	FALSE	TRUE	FALSE	FALSE
13	11	20090617	1	0			2009	TRUE	FALSE	FALSE	FALSE
57	11	20090621	0				2009	FALSE	FALSE	FALSE	FALSE
59	11	20090802	0				2009	FALSE	FALSE	FALSE	FALSE
54	11	20090912	3				2009	FALSE	FALSE	TRUE	FALSE
71	11	20091019	1	0			2009	TRUE	FALSE	FALSE	FALSE
33	11	20091130	1	50			2009	TRUE	FALSE	FALSE	FALSE
50	11	20091208	1	0			2009	TRUE	FALSE	FALSE	FALSE
6	11	20091230	0	1			2009	TRUE	FALSE	FALSE	FALSE

Source code = 1

Count of AGE				
Row Labels	2009	2010	2011	Grand Total
FALSE	160	417	264	841
TRUE	447	210	531	1188
Grand Total	607	627	795	2029

Source code = 2

Count of AGE				
Row Labels	2009	2010	2011	Grand Total
FALSE	601	623	793	2017
TRUE	6	4	2	12
Grand Total	607	627	795	2029

Source code = 3

Count of AGE				
Row Labels	2009	2010	2011	Grand Total
FALSE	606	620	786	2012
TRUE	1	7	9	17
Grand Total	607	627	795	2029

Source code = 12

Count of AGE				
Row Labels	2009	2010	2011	Grand Total
FALSE	583	590	756	1929
TRUE	24	37	39	100
Grand Total	607	627	795	2029

Table 9. 8 Calculating no. of sources per registration: Example of an Excel Export File

Table 9. 9 shows an example of a table of results. This is a very useful check on reporting from the different sources, and suggests where case finding might have been incomplete at certain periods.

UASIN GISHU CASES: Number of sources

Source	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Total
00 DEATH CERTIFICATE	9	34	19	29	48	57	28	85	97	97	114	48	0	0	665
01 MR&TH	176	337	318	255	198	273	125	253	214	418	548	494	181	610	4400
02 ELDORET	11	14	17	16	22	12	12	10	2	4	7	3	2	1	133
03 PACIFICA	3	4	0	0	2	0	0	0	0	0	2	0	0	0	11
04 U.G MEMORIAL*	35	27	20	12	12	1	12	0	0	0	0	0	0	0	119
05 Kenyatta National	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
06 M.P. Shah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
07 NAIROBI	1	2	0	0	0	0	1	1	1	0	1	0	1	0	8
10 Private Clinics	13	9	2	2	15	15	13	10	5	7	9	6	1	2	109
11 Private labs	0	0	0	2	0	0	5	21	1	2	3	3	0	1	38
12 HOSPICE	104	76	33	38	25	39	0	0	0	0	0	0	0	0	315
13 ELGON VIEW HOSPITAL	0	3	6	8	10	3	6	3	0	0	2	2	0	0	43
14 ITEN	0	0	2	0	1	0	0	0	0	0	0	0	0	1	4
15 KAPSABET	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
16 KITALE/MT.ELGON	0	0	3	0	0	3	15	0	0	0	0	0	0	2	23
17 KAPSOWAR	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2
18 KAPEGURI	1	0	0	0	0	0	0	0	0	0	0	0	0	3	4
25 PLATEAU	1	0	0	0	3	0	0	0	0	0	1	2	0	0	7
50 HISTOLOGY LABS	58	217	123	159	159	213	126	210	59	266	328	377	424	751	3470
51 HAEMATOLOGY	5	14	33	32	67	78	46	26	50	11	12	40	170	113	697
60 RADIOLOGY	21	11	37	15	11	0	0	0	0	0	0	0	0	0	95
all sources	438	750	614	569	574	694	393	623	429	805	1030	978	779	1488	10164
cases	353	487	409	343	331	366	233	350	241	407	500	518	589	845	5972
average no. of sources per case	1.24	1.54	1.50	1.66	1.73	1.90	1.69	1.78	1.78	1.98	2.06	1.89	1.32	1.76	1.70

Table 9. 9 Analysis of the number of notifications from each source, by year (Eldoret Cancer Registry)

9.2.1.2 Comparison of incidence rates in different populations

The possibility of incomplete registration can be investigated by comparing observed incidence rates with expected values, based on those observed in registries in the same region, or estimated for the country in the latest edition of GLOBOCAN. The assumption is that the incidence rates for specific cancers should be rather similar to those from elsewhere in the same region (or country).

A standard Table (Data Quality Indicators) similar to that used by the editors of CI5 (Table 9.10) is planned for the "Table builder" option of CanReg 5 for this purpose.

This table presents the age-standardized incidence rates (and their standard errors) for 21 sites (and the total for all sites) in males and females, along with the ratio of the observed value to the expected value (O/E).

If the observed age standardized rate is significantly different from the expected value for the corresponding country or region, the O/E is shown in bold and flagged with a greater-than symbol (>) if the value is higher than expected or a less-than symbol (<) if the value is lower than expected.

EREWHON (2003–2007)

MALE

SITE	Cases	ASR (se)	O/E	MV(%)	DCO(%)	M/I(%)	ICD-10
Lip, oral cavity and pharynx	6968	14.5 (0.18)	0.98	91.2	7.2	38.9	C00–14
Oesophagus	3293	6.3 (0.11)	0.95	83.4	14.7	75.4	C15
Stomach	7481	**12.5 (0.15) <**	**0.83**	76.9	19.5	61.2	C16
Colon, rectum and anus	27365	47.3 (0.30)	1.09	85.6	13.1	38.4	C18–21
Liver	4185	**7.4 (0.12) >**	**1.26**	58.0	37.3	82.9	C22
Pancreas	5339	9.3 (0.13)	1.03	56.0	35.7	91.4	C25
Larynx	2280	**4.4 (0.10) <**	**0.88**	87.0	12.0	39.3	C32
Lung (incl. trachea)	20052	**35.2 (0.26) <**	**0.67**	**71.7 <**	25.1	83.0	C33–34
Melanoma of skin	5703	**11.5 (0.16) >**	**1.32**	94.7	3.7	20.8	C43
Prostate	46799	**79.2 (0.38) >**	**1.33**	86.3	11.1	**17.8 <**	C61
Testis	2922	8.5 (0.17)	1.07	**92.9 <**	2.1	3.6	C62
Kidney etc.	7345	13.4 (0.17)	0.92	84.8	13.1	41.4	C64–66
Bladder	12303	20.5 (0.19)	0.98	91.2	7.6	**16.5 <**	C67
Brain, central nervous system	2777	6.3 (0.13)	1.04	73.7	21.6	70.5	C70–72
Thyroid	1493	**3.3 (0.09) >**	**1.70**	93.5	5.0	**14.8 <**	C73
Lymphoma	7747	15.4 (0.19)	1.02	**81.7 <**	16.3	43.5	C81–88,C90
Leukaemia	4619	**9.5 (0.17) <**	**0.88**	70.3	29.0	62.0	C91–95
Ill–defined (2.2% of total)	3941	6.8 (0.11)	0.87	50.7	43.4	132.9	C76–80
All sites but non–melanoma skin	179172	324.2 (0.81)	1.02	81.6	15.8	43.7	C00–96bC44

FEMALE

SITE	Cases	ASR (se)	O/E	MV(%)	DCO(%)	M/I(%)	ICD-10
Lip, oral cavity and pharynx	2076	3.7 (0.09)	1.04	89.5	8.8	33.4	C00–14
Oesophagus	741	1.1 (0.04)	0.84	76.2	22.0	75.6	C15
Stomach	6017	6.9 (0.11)	0.90	70.7	25.2	67.0	C16
Colon, rectum and anus	22635	28.1 (0.22)	1.02	80.4	18.1	41.7	C18–21
Liver	1621	2.1 (0.06)	1.02	48.5	46.7	93.8	C22
Pancreas	5433	6.2 (0.10)	1.01	44.8	47.5	93.6	C25
Larynx	274	0.5 (0.03)	0.85	84.7	14.2	40.5	C32
Lung (incl. trachea)	8461	13.1 (0.16)	0.95	72.2	24.3	78.4	C33–34
Melanoma of skin	5795	**11.3 (0.16) >**	**1.27**	94.8	3.8	15.3	C43
Breast	48551	**82.6 (0.41) >**	**1.15**	**86.9 <**	11.2	26.9	C50
Cervix uteri	3523	**7.1 (0.13) <**	**0.76**	89.5	8.4	31.8	C53
O&U part of uterus	8689	13.0 (0.15)	1.03	**88.8 <**	10.0	23.3	C54–55
Ovary	6672	10.4 (0.14)	1.01	**75.2 <**	21.8	64.9	C56
Kidney etc.	4702	6.6 (0.11)	0.89	79.9	17.6	40.9	C64–66
Bladder	4315	5.2 (0.09)	1.04	86.3	12.3	**25.0 <**	C67
Brain, central nervous system	2405	4.6 (0.11)	1.02	66.9	28.2	67.2	C70–72
Thyroid	3778	**8.4 (0.15) >**	**1.68**	94.3	4.5	**10.7 <**	C73
Lymphoma	7156	11.3 (0.16)	1.04	**77.8 <**	20.2	46.7	C81–88,C90
Leukaemia	3743	6.1 (0.13)	0.90	62.7	36.8	68.6	C91–95
Ill–defined (2.7% of total)	4228	4.5 (0.08)	0.89	38.9	55.1	114.0	C76–80
All sites but non–melanoma skin	158408	**243.8 (0.71) >**	**1.05**	79.4	18.1	44.0	C00–96bC44

Data compared with those from seven registries in the same region/country.
Significantly lower (<) or higher (>) values are shown in bold.

Table 9. 10 Data Quality Indicators: Standard Table

9.2.1.3 Shape of age-specific curves

CanReg 5 produces a set of graphs showing age-specific incidence for 12 cancer sites (one curve for each sex) – option "Age-specific rates for major/most diagnosis groups (semi-logarithmic)" in Table builder (Figure 9.2).

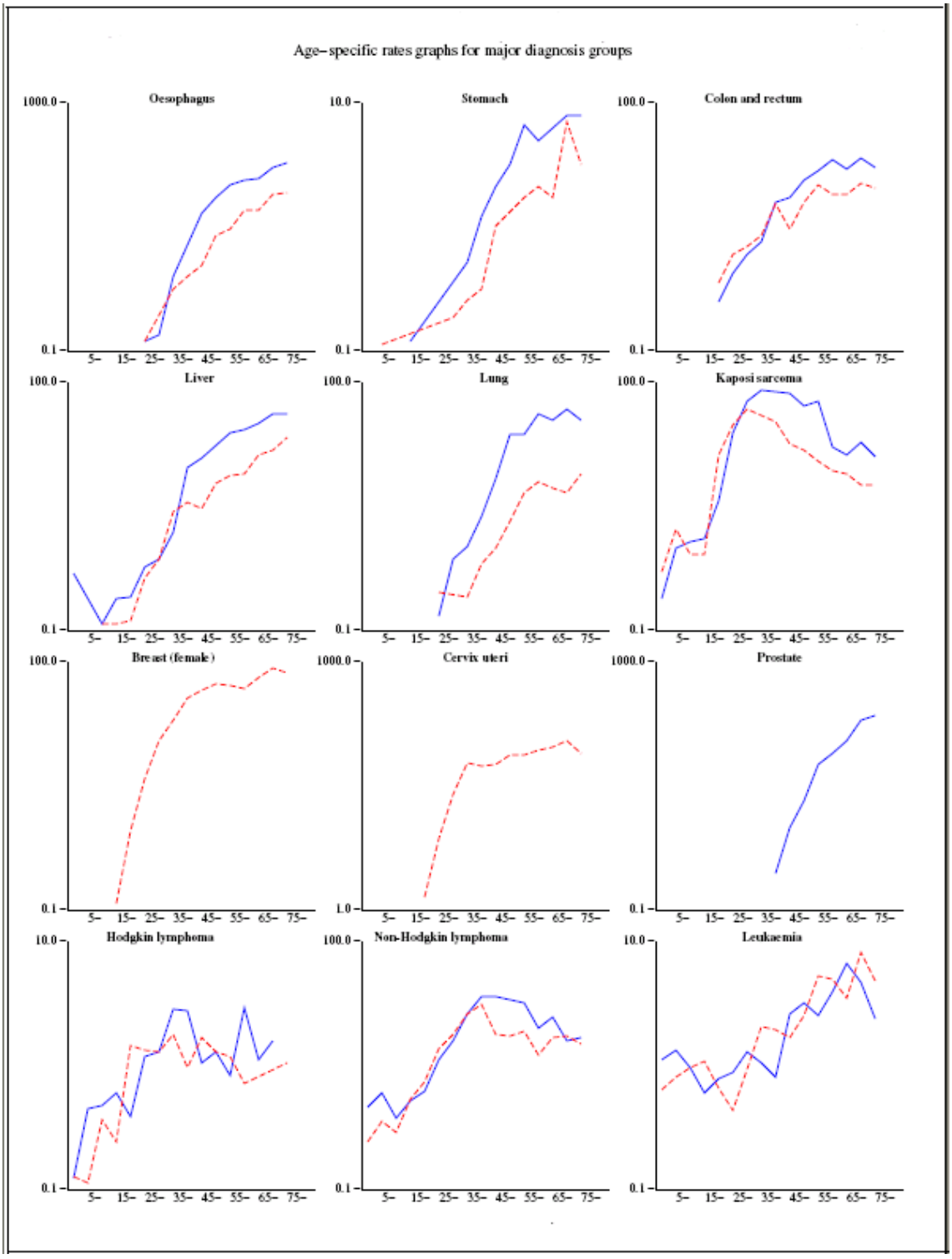


Fig 9. 2 Age-specific rates graphs for major diagnosis groups

The curves can be examined in order to detect abnormal fluctuations in the anticipated patterns, including any fall-off in the rate of increase in incidence in older subjects (which may be indicative of under ascertainment within these groups (although there can also be other explanations).

9.2.1.3 Incidence rates of childhood cancers

With respect to childhood cancer, the incidence rates (for all types combined) in the childhood age groups (0-4. 5-9, and 10-14) show much less variability than in adults, The possibility of under-enumeration (or duplicate registrations) in this age range can be investigated by comparing the observed age-specific rates in the childhood age range with an "expected" range of values.

The lowest and highest deciles of incidence rates of childhood cancer in the CI5 Volume XI data are shown in Table 9. 11.

The lowest and highest deciles of incidence rates (per 100 000) of childhood cancer in Volume X

Age group (years)	Boys		Girls	
	Lowest	Highest	Lowest	Highest
0–4	< 12.6	> 26.4	< 12.1	> 23.7
5–9	< 8.9	> 17.9	< 7.0	> 13.0
10–14	< 9.0	> 17.2	< 8.2	> 16.0

Table 9. 11 The lowest and highest deciles of incidence rates (per million) of childhood cancer in Volume XI

9.2.2. Mortality:Incidence Ratios

The Mortality:Incidence (M:I) ratio is an important indicator of completeness for cancer registries. It compares the number of deaths, obtained from a source independent of the registry (usually, the vital statistics system), and the number of new cases of a specific cancer registered, in the same period of time.

This method cannot be used where there is no comprehensive death registration, or when cause of death is missing or inaccurate on death certificates, which is the situation in almost all countries in Africa.

For the very few countries with reasonable death registration, the M:I ratios can be compared to standard values from the same region, testing for significant differences (see Table 9.10). When the quality of the mortality data is good and incidence and survival are in steady state, the M:I ratio is approximately 1 minus the 5-year survival probability (Asadzadeh Vostakolaei et al, 2011).

M:I ratios that are higher than expected raise suspicion of incompleteness (i.e. incident cancers missed by the registry), especially if the values are high for several different sites. However, under- or over reporting of tumours on the death certificates distorts this relationship.

9.2.3. Number of sources/notifications per case

The reason for using as many sources of information as possible on cancer cases, is that it reduces the likelihood of missing cases. Multiple sources reporting thus increase the completeness of the registry data.

When examining the number of reports from different sources, by year of reporting (see section 9.2.1.1) it is simple to compare the total number of sources reporting cases in one year, with the number of cases registered (See Table 9. 9). The ratio is the average number of sources per case.

In the example shown (Table 9.9) it is clear that the average is low (1.3) in one year (2010), compared with the average for the whole 14 year period (1.7 sources per case) due to poor reporting from two sources (MR& TH and Death Certificates) in that year.

9.2.4 Independent case ascertainment

Two methods can be used:

- Re-screening the sources that had been used by the registry, to detect any case missed during the registration process (Case finding audits);
- Using independent sources of cancer cases (which have not been used by the registry), and comparison of the registry database with them.

9.2.4.1 Case finding audit

Here, the idea is to go back to one or more of the registries data sources, and do an independent case finding (and abstracting) exercise. NAACCR suggests that each source ("reporting facility") should be routinely audited at least once every 3 years (Hofferkamp, 2008). Audits also should be conducted when there is a documented decline in case reports from a facility (compared with the numbers of reports in the previous year's – see Section 9.2.1.1).

Records of cancer cases identified during the audit are enumerated and matched against the registry's files. Unmatched cases are followed back to verify whether they meet the reportability criteria (Section 3.2). The percentage of cases actually missed that *should* have been reported is calculated.

Most such studies focus on hospital sources. They thus provide an estimate of the completeness of reporting for those sources, not a true estimate of completeness for the whole registry (which is using multi-source reporting).

9.2.4.2 Using an independent source

Here, we need to find a list of cancer cases that have been compiled *independently* of the cancer registry's case-finding procedures.

Comparing this list of cases with the registry database is a particularly useful method of evaluating completeness.

It requires record linkage between the cancer registry database and the independent case series, to estimate the numbers of cases in the latter "missed" by the registry.

Record linkage can be done using the "REC-LINK" software, originally developed by IARC[6] or other more recently developed free record linkage software such as LinkPlus:

http://www.cdc.gov/cancer/npcr/tools/registryplus/lp.htm

or FRIL :

http://fril.sourceforge.net/

The proportion of eligible patients who are already registered is a direct and quantitative estimate of completeness. What independent sources of cancer cases might exist?

The most usual sources in Africa are:

- Cases recruited into local clinical trials
- Cases recruited into special studies (e.g. hospital case control studies)

[6] REC-LINK is available from the AFCRN secretariat

- Cases recorded in databases by individual clinicians
- Cancer deaths from special community studies (e.g. verbal autopsy studies)

The text box below describes a study of completeness of registration in the Kampala Cancer Registry, using an independent data source (a case-control study of HIV and cancer), which recruited cancer cases from the main hospital, using project staff who worked independently of the cancer registry (Parkin et al, 2001).

THE KAMPALA COMPLETENESS STUDY

Record linkage and verification

A file was prepared of cases of cancer enrolled into the HIV Cancer study, with a date of diagnosis during 1994-1996 and resident in Kyadondo County. This file was linked to the master file of the cancer registry in June 1998 (approximately 18 months after the recorded diagnosis of the last case).

Record linkage was performed using the program REC-LINK. This provides a probabilistic matching of files based upon selected variables; in the current study, we used family name, first name, age, sex, and tribe. Each variable is assigned a score from 1 to 5 for reliability (reflecting the probability that individuals who are truly identical will have identical records with respect to the variable) and discriminating power (reflecting the probability that identical records with respect to the variable truly belong to the same individual). The product of the reliability score and the discriminating powers score gives a weighting factor for that variable.

When two records are compared, the two values for each variable are compared to arrive at a Similarity score. For each variable, its weighting is multiplied by this similarity score, and the sum (expressed as a percentage) gives the final probability that these two records belong to the same person.

Possible linkages (with percentage scores between 70 and 90) were checked "manually", using additional recorded variables (including maiden name, address, diagnosis, hospital, laboratory number) to make the final decision.

Cases that could not be linked to the cancer registry were examined in more detail. The study questionnaires were reviewed and, for many of the subjects, so too were many cases, their inclusion in the data set for record linkage was an error. This was the result of mistakes in recording of information, or in data coding and data entry, especially with respect to date of diagnosis (several prevalent cancer cases had been enrolled into the study), diagnosis (some enrolled cases had proved not to have cancer when the full range of diagnostic tests was completed), and address. These cases were deleted from the file.

Estimation of completeness

Case-finding for the cancer registry was performed independently of recruitment into the HIV Cancer study and completeness of registration was therefore estimated from the percentage of eligible cases in the HIV Cancer study data set which had been included in the registry database. Ninety-five percent confidence limits were calculated by the "exact" method.

To investigate the independent contribution of different patient variables to completeness of ascertainment, we fitted a logistic regression model to the data using STATA. The outcome variable was detection (or not) of cases by the registry, and the explanatory variables were sex, age (continuous and grouped), year, basis of diagnosis (microscopic or not), and diagnosis. The initial age groups used were 15-29, 30-39, 40-49, 50- 59, and 60-69. Diagnoses were initially considered as 23 cancers, and then regrouped into seven categories (Kaposi's sarcoma, cervix cancer, oesophagus, liver, breast, eye, and all other).

9.2.5 Capture-Recapture Methods

Capture-recapture methods are possible when cancer registries use two or more different types of source for case finding. For registries using CanReg (version 4 or 5), each cancer case registered in the database may have several sources (where the information on the case was found). For a capture-recapture analysis, the different sources need to be regrouped into two or three reasonably independent categories:

Source 1: Hospital Source (hospital records departments/ radiotherapy/ oncology etc)
Source 2: Pathology Labs (include cytology, haematology)
Source 3: Death Certificates

For capture-recapture estimates of completeness to be accurate, identification (capture) of a case by one type of source should be independent of the other(s). In practice, this is a bit unlikely. For example, cases of cancer admitted to hospital might be more likely to be found in the pathology laboratory records labs than cases that were not. Or patients who die (and get a Death Certificate) may be less likely to be found from pathology (if they died soon after reaching hospital), than patients who do not die.

Although these limitations may cause estimates of completeness to be a bit inaccurate, in practice they are not too bad (Parkin et al, 1994).

Estimate with TWO types of source

When only two types of source are available (usually this will be hospital and pathology sources), cases registered in a given period (usually one year) can be classified as being registered from information obtained from one, or the other, or both, as in Table 9.12.

		SOURCE A	
		PRESENT	ABSENT
SOURCE B	PRESENT	***a***	***b***
	ABSENT	***c***	***d***

Table 9. 12 Estimate of Completeness by Capture-Recapture using TWO types of source

The number of cases from neither source (*d*) represent cases "missed" by case finding.

d is estimated as **bc/a**

and completeness will be $\frac{(a+b+c)}{(a+b+c+d)}$

Estimate with THREE types of source

If registrations can be classified by all three types of source, then 3 estimates of completeness can be made, and the true values should be somewhere between the highest and lowest values. The method is to use the above "paired" estimate, using, for each source, the combination of the other two as the other half of the pair. Using the example of Hospital (HOSP), Pathology (PATH) and Death Certificate (D.C.) as three sources, there are 7 possible combinations of sources (Fig 9. 3), and three sets of paired tables can be constructed (Table 9. 13).

For each table, we estimate the number of cases d that appear in none of the three sources.

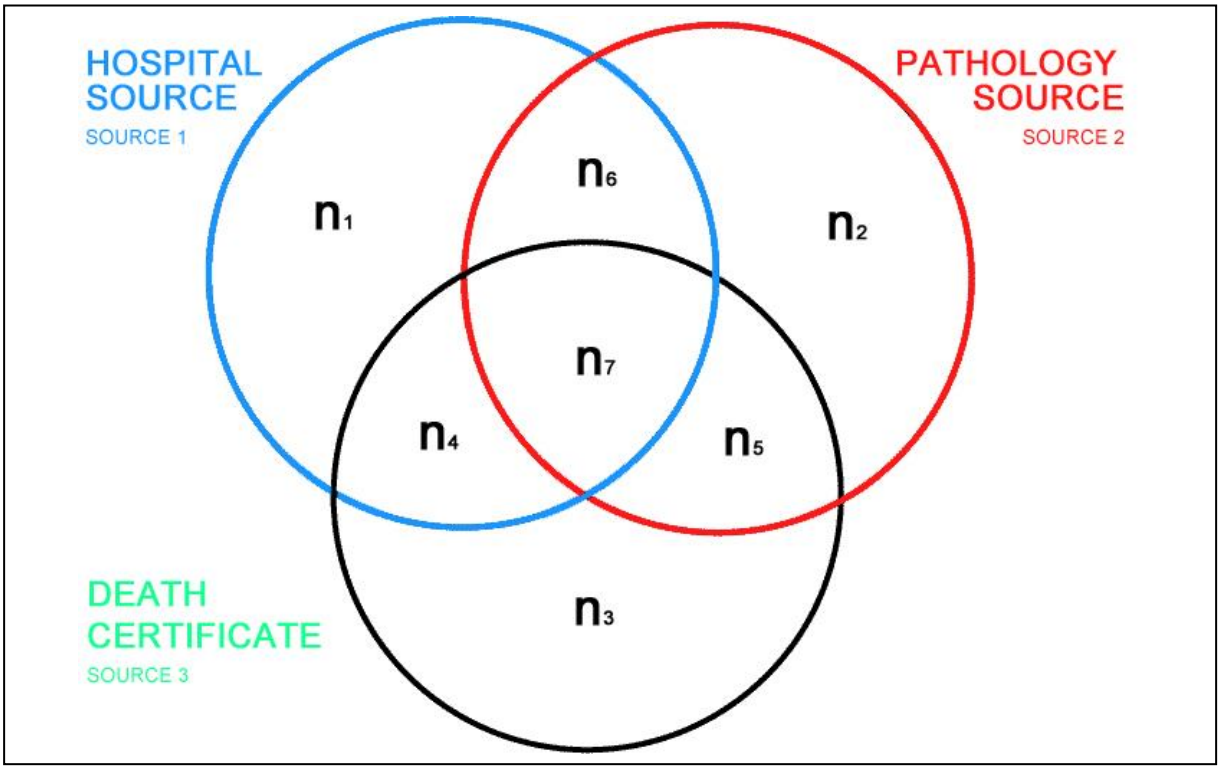


Fig 9. 3 Possible Combinations of THREE Sources

		HOSPITAL	
		PRESENT	ABSENT
PATH. and/or D.C.	PRESENT	$n_6+n_7+n_4$	$n_2+n_5+n_3$
	ABSENT	n_1	***d***

		PATHOLOGY	
		PRESENT	ABSENT
HOSP and/or D.C.	PRESENT	$n_6+n_7+n_5$	$n_1+n_4+n_3$
	ABSENT	n_2	***d***

		Death Certificate	
		PRESENT	ABSENT
PATH and/or HOSP.	PRESENT	$n_4+n_7+n_5$	$n_1+n_6+n_2$
	ABSENT	n_3	***d***

Table 9. 13 Estimate of Completeness by Capture-Recapture using THREE types of sources

Differences between the estimates are due to the interdependence of pairs of sources (clinical records, pathology, death certificates).

Various more sophisticated methods are available for dealing with the problem of dependency between sources. Larsen et al (2004) identified the degree of dependence between pairs of sources, then grouped those sources with the most dependence, before estimating the missing cases in a two way method with a third source. When the sources are all dependent, this approach cannot work, and log-linear modelling is needed (Parkin & Bray, 2009).

9.2.6 Death Certificate Methods

Using death certificates as a source of information can provide a very useful method of evaluating completeness. If the registry is finding many new cases via death certificates, it is certain that registration is incomplete (since some patients who do NOT die will have been missed by case finding).

To estimate completeness, we need to know the number of cases identified via a death certificate, that would otherwise never have been found; these are "death certificate initiated" (DCI) cases. These are not the same as Death Certificate Only (DCO) cases – the latter are only a fraction of the cases which are first identified from a death certificate (those that cannot be traced by follow-back to the source of the death certificate (see Section 6. DEATH CERTIFICATE NOTIFICATIONS, page 35)).

Two methods are described in Parkin & Bray (2009) for estimating completeness using death certificates; the DCI:MI method and the Flow method. Neither are applicable if comprehensive death registration by cause (vital statistics system) is not available. It is however possible to use two-source capture-recapture, as described above, even in circumstances where death registration is incomplete or inaccurate. To use this method, the registry must record all the sources of notification for each case – i.e., was it found from a death certificate, another source, or both? Cases can then be tabulated as in Fig 9.4

Fig 9.4

		DEATH CERTIFICATES	
		PRESENT	ABSENT
OTHER SOURCES	PRESENT	a	b
	ABSENT	c	d

The box labelled (a) shows the number of cases that are found in the registry database (from other sources) AND via death certificates. Box (b) shows the number of cases found in the registry from other sources but not from a death certificate. The box (c) the cases that are DCI (found from death certificates, but NOT in the registry database). Lastly, box (d) is the value that represents the missed cases – either alive (d_1) or dead, but not identifiable via a death certificate (d_2) – and is the value that is needed to estimate completeness.

Completeness is calculated as:

$$Completeness = \frac{a+b+c}{a+b+c+d}$$

Where d is given by :

$$d = \frac{b \times c}{a}$$

An important assumption lies behind the use of this method to estimate *d* (Lincoln-Peterson estimator): that the chance of inclusion in one source (death certificates) is independent of inclusion in the other (other sources of information). Or, put another way, that case fatality will be the same in the cases found from other (non-DC) sources, and those that were not.

HOWEVER, the estimate is not affected by the completeness or accuracy of death registration (assuming that any incompleteness or inaccuracy is independent of the cause of death (cancer or not)), nor by the case fatality (the proportion of cases that die).

The competitions on limited resources in Africa due to other health problems that include HIV/AIDS make cancer registration sometimes a dream hard to realise. For those registries managed in whatever form; do not give up, keep going and remember that with imagination, patience, perseverance, dedication and willingness to solve local problems; establish population-based cancer registration is possible.

REFERENCES

- Asadzadeh Vostakolaei F, Karim-Kos HE, Janssen-Heijnen ML, Visser O, Verbeek AL, Kiemeney LA. The validity of the mortality to incidence ratio as a proxy for site-specific cancer survival. Eur J Public Health. 2011 ;21(5):573-7
- Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009 ;45(5):747-55.
- Curado MP, Edwards B, Shin HR Storm H, Ferlay J, Heanue M and Boyle P. Cancer Incidence in Five Continents - Vol.IX. (IARC Scientific Publication No 160). Lyon, IARC, 2007
- Esteban D, Whelan S, Laudico A, Parkin DM (1995). Manual for cancer registry personnel. IARC Technical Report No.10. IARC, Lyon.
- Fritz A, Percy C, Jack A,Shanmugaratnam K, Sobin L,Parkin DM, Whelan S. (2000)WHO International Classification of Disease for Oncology. Third Edition, Geneva.
- Hilsenbeck SG, Glaaefke GS, Feigl P, Lane WW, Golenzer H, Ames C, et al. Quality Control for Cancer Registries. Bethesda, MD: U.S. Department of Health and Human Services, Public Health Service, National nstitutes of Health; 1985. Third printing, U.S. Government Printing Office No. 1987-793-588; May 1987.
- Hofferkamp, J (Ed). Standards for Cancer Registries Volume III: Standards for Completeness, Quality, Analysis, Management, Security and Confidentiality of Data. Springfield (IL): North American Association of Central Cancer Registries, August 2008.
- Jensen OM, Parkin DM, Maclennan R, Muir CS,Skeet RG (1991)Cancer registration principles and methods. IARC Scientific publications No 95. IARC, Lyon.
- Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, Moller B. Data quality at the Cancer Registry Norway: an overview of comparability, completness, validity and timeliness. Eur J Cancer. 2009; 45(7): 1218-31
- Louange E. (2008) Seychelles Reporting and Coding Manual, Ministry of Health .No 1.
- Parkin DM, Bray F. Evaluation of data quality in the cancer registry:principles and methods Part II. Completeness. Eur J Cancer. 2009;45(5):756-64
- Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH & Whelan SL. Comparability and Quality Control in Cancer Registration. IARC Technical Report No. 19. Lyon, IARC, 1994
- Parkin DM, Wabinga H, Nambooze S. Completeness in an African cancer registry. *Cancer Causes Control*. 2001; 12, 147-152.
- Sobin LH, Gospodarowicz MK, Wittekind C. ,TNM Classification of Malignant Tumours 7th Edition. Union International Cancer Control
- Somdyala N, (2008) Procedure and Guidelines for Data Management, PROMEC Cancer Registry, South AfricaOkondo F, Draft Standard Operating Procedures (2014) Gulu Cancer Registry, Uganda
- World Health Organisation. International Classification of Diseases and Related Health Problems. 10th edition, WHO, Geneva, 2010.
- Wu XC, McLaughlin C, Lake A, Firth R, Leonfellner S, Cormier M, Cardinez C, Kosary C, Roney D & Howe H. Cancer in North America, 2000-2004. Volume I, Incidence. A Publication of the North American Association of Central Cancer Registries, Inc.(NAACCR) May 2007

APPENDIX 1. MULTIPLE PRIMARY CANCERS - RULES FOR REPORTING INCIDENCE AND SURVIVAL

1. The recognition of the existence of two or more primary cancers does not depend on time.

2. A primary cancer is one that originates in a primary site or tissue and is not an extension, nor a recurrence, nor a metastasis.

3. Only one tumour shall be recognised as arising in an organ or pair of organs or tissue.

Some groups of codes are considered to be a single organ for the purposes of defining multiple tumours. These topography code groups are shown in Appendix Table 1.

Multifocal tumours – that is, discrete masses apparently not in continuity with other primary cancers originating in the same primary site or tissue, for example bladder – are counted as a single cancer.

4. Rule 3 does not apply in two circumstances:
   4.1 Systemic (or multicentric) cancers potentially involving many different organs are only counted once in any individual. These are Kaposi sarcoma (group 15 in Table 2) and tumours of the haematopoietic system (groups 8-14 in Appendix Table 2).
   4.2 Neoplasms of different morphology should be regarded as multiple cancers (even if they are diagnosed simultaneously in the same site).

If the morphological diagnoses fall into one category in Appendix Table 2, and arise in the same primary site, they are considered to be the same morphology for the purpose of counting multiple primaries. If the morphological diagnoses fall into two or more of the categories in Appendix Table 2, even if they concern the same site, the morphology is considered to be different, and two or more cases should be counted.

Single tumours containing several different histologies which fall into one histological group in Appendix Table 2 are registered as a single case, using the numerically highest ICD-O morphology code.

If, however, one morphology is not specific (groups (5), (14) and (17)) and a specific morphology is available, the case should be reported with the specific histology and the non-specific diagnosis should be ignored.

Table 1. Groups of topography codes considered a single site in the definition of multiple cancers

ICD-O-2/3 site code	Label	If diagnosed at different times, code first diagnosis. If diagnosed at the same time use codes given below.
C01	Base of tongue	
C02	Other and unspecified parts of tongue	C02.9
C00	Lip	
C03	Gum	
C04	Floor of mouth	
C05	Palate	
C06	Other and unspecified parts of mouth	C06.9
C09	Tonsil	
C10	Oropharynx	
C12	Pyriform sinus	
C13	Hypopharynx	
C14	Other and ill-defined sites in lip, oral cavity and pharynx	C14.0
C19	Rectosigmoid junction	
C20	Rectum	C20.9
C23	Gallbladder	
C24	Other and unspecified parts of biliary tract	C24.9
C33	Trachea	
C34	Bronchus and lung	C34.9
C40	Bones, joints and articular cartilage of limbs	
C41	Bones, joints and articular cartilage of other and unspecified sites	C41.9
C65	Renal pelvis	
C66	Ureter	
C67	Bladder	
C68	Other and unspecified urinary organs	C68.9

Appendix Table 1 Groups of topography codes considered a single site in the definition of multiple cancers

Table 2. Groups of malignant neoplasms considered to be histologically 'different' for the purpose of defining multiple tumours (adapted from Berg JW. Morphologic classification of human cancer. In: Schottenfeld D & Fraumeni JF Jr. *Cancer Epidemiology and Prevention*, 2nd edition, Chapter 3 of Section 1: Basic Concepts. Oxford, New York, Oxford University Press, pp. 28-44, 1996).

Group	
Carcinomas	
1. Squamous and transitional cell carcinoma	8051-8084, 8120-8131
2. Basal cell carcinomas	8090-8110
3. Adenocarcinomas	8140-8149, 8160-8162, 8190-8221, 8260-8337, 8350-8551, 8570-8576, 8940-8941
4. Other specific carcinomas	8030-8046, 8150-8157, 8170-8180, 8230-8255, 8340-8347, 8560-8562, 8580-8671
(5) Unspecified carcinomas (NOS)	8010-8015, 8020-8022, 8050
6. *Sarcomas* and soft tissue tumours	8680-8713, 8800-8921, 8990-8991, 9040-9044, 9120-9125, 9130-9136, 9141-9252, 9370-9373, 9540-9582
7. ***Mesothelioma***	9050-9055
Tumours of haematopoietic and lymphoid tissues	
8. Myeloid	9840, 9861-9931, 9945-9946, 9950, 9961-9964, 9980-9987
9. B-cell neoplasms	9670-9699, 9728, 9731-9734, 9761-9767, 9769, 9823-9826, 9833, 9836, 9940
10. T-cell and NK-cell neoplasms	9700-9719, 9729, 9768, 9827-9831, 9834, 9837, 9948
11. Hodgkin lymphoma	9650-9667
12. Mast-cell Tumours	9740-9742
13. Histiocytes and Accessory Lymphoid cells	9750-9758
(14) Unspecified types	9590-9591, 9596, 9727, 9760, 9800-9801, 9805, 9820, 9832, 9835, 9860, 9960, 9970, 9975, 9989
15. ***Kaposi sarcoma***	9140
16. ***Other specified*** types of cancer	8720-8790, 8930-8936, 8950-8983, 9000-9030, 9060-9110, 9260-9365, 9380- 9539
(17) ***Unspecified*** types of cancer	8000-8005

Appendix Table 2 Groups of malignant neoplasms considered to be histologically 'different' for the purpose of defining multiple tumours

RECOMMENDATIONS FOR RECORDING

A. Two tumours of different laterality, but of the same morphology, diagnosed in paired organs (e.g. breast) should be registered separately unless stated to have originated from a single primary. Exceptions to this rule are:
    a. Tumours of the ovary (of the same morphology)
    b. Wilm's tumour (nephroblastoma) of the kidney.
    c. Retinoblastoma

   which should be recorded as a single bilateral registration when they occur on both sides.

   ***Reminder*:** tumours in paired organs of completely different histology should be registered separately.

B. Cancers which occur in any 4th character subcategory of colon (C18) and skin (C44) should be registered as multiple primary cancers.

APPENDIX 2. CANCER CHEMOTHERAPEUTIC AGENTS

Chemotherapeutic Agent	Brand Name(S)
5 AZACYTIDINE	Vidaza
5 FLUOROURACIL	Adrucil;fluoroblastin;5 FU
ABRAXANE	Abraxane
ACTINOMYCIN	Cosmegen; dactinomycin
AMSACRINE	M-AMSA;4'-(9 acridinyl aminomethane-sulfon-m-anisodide)
ASPARAGINASE	Crisantaspase (Erwinase)
AZACITIDINE	Vidaza
BENDAMUSTINE	Levact
BLEOMYCIN	Blenoxane
BUSULFAN	Myleran
CABAZITAXEL	Jevtana
CAPECITABINE	Xeloda
CARBOPLATIN	Paraplatin
CARMUSTINE	BCNU;bichlorethylnitrosourea
CHLORAMBUCIL	Leukeran
CISPLATIN	Cis-diammine dichloroplatinum DDP;platinol;platamine
CLADRIBINE	Leustat, LITAK
CLOFARABINE	Evoltra
CRISANTASPASE	Erwinase, asparaginase or L-asparaginase
CYCLOPHOSPHAMIDE	Cytoxan; cyclophar; endoxan
CYTARABINE	Cytosine arabinoside;arabinosyl cystosine;cytosar-U; ara-C
DACARBAZINE	DTIC; dimethyltriazine imidazole carboxamide
DACTINOMYCIN	Cosmegen Lyovac
DAUNORUBICIN	Daunomycin;cerubidin
DOCETAXEL	Taxotere
DOXORUBICIN	Adriamycin
DOXORUBIN	Adriamycin;hydroxyl daunorubicin
EPIRUBICIN	Pharmorubicin
ETOPOSIDE	VP-16, Etopophos, Vepesid
FLUDARABINE	Fludara
FLUOROURACIL	5FU
GEMCITABINE	Gemzar
GLIADEL IMPLANTS	Gliadel implants (carmustine)
HEXAMETHYL-MELAMINE	HMM
HYDROXYCARBAMIDE	(Hydrea, hydroxyurea)
HYDROXYUREA	Hyrea
IDARUBICIN	Zavedos
IFOSFAMIDE	Holoxan
IRINOTECAN	Campto
LEUCOVORIN	(folinic acid)
LIPOSOMAL DAUNORUBICIN	DaunoXome

LIPOSOMAL DOXORUBICIN	Caelyx, Myocet
LOMUSTINE	CCNU;cyclophexylchlorethyl nitosourea;CeeNU
MELPHALAN	Alkeran;phenylalanine mustard;L-PAM;L- Sarcolysin
MERCAPTOPURINE	6-MP;purinethol
MESNA	Uromitexan
METHOTREXATE	Amethopterin; MTX, matrex; mexate emthexate
MITHRAMYCIN	Mithracin
MITOMYCIN	Mitomycin C Kyowa
MITOMYCIN C	Mutamycin
MITOTANE	op'-DDP;lysodren
MITOXANTRONE	Novantrone;DHAD
OXALIPLATIN	eloxatin
PACLITAXEL	Taxol
PEMETREXED	Alimta
PENTOSTATIN	Nipent
PROCARBAZINE	Matulane;methylhydrazine
RALTITREXED	Tomudex
RASBURICASE	Fasturtec
SEMUSTINE	Methyl-CCNU;MeCCNU;chloroethyl methylcyclohexyl nitrosourea
STREPTOZOZIN	Streptozotocin
TAXOL	Paclitaxel (taxol)
TEGAFUR-URACIL	Uftoral
TEMOZOLOMIDE	Temodal
THIOGUANINE	6-TG;lanvis
THIOTEPA	Thio-TEPA;triethylenethiophosphoramide
TIOGUANINE	lanvis
TOPOTECAN	Hycamtin
TRABECTEDIN	Yondelis
TREOSULFAN	Ovastat, Treosulfan Medac, Treosulfan
VINBLASTINE	Velban
VINCRISTINE	Oncovin
VINDESINE	Eldisine
VINORELBINE	Navelbine
VP-16	Etopside;VP-16-213 epipodophyllotoxin;vepesid

APPENDIX 3. CONFIDENTIALITY AGREEMENT FORM TEMPLATE

[YOUR LOGO HERE]

CONFIDENTIALITY AGREEMENT

I, _____, agree with the following statements:

I have read and understood [name of registry]'s Privacy Policy.
I understand that I may come in contact with confidential information during my time at **[name of registry].** As part of the condition of my work with **[name of registry]** I hereby undertake to keep in strict confidence any information regarding any patient (alive or dead), client, employee or any other organization that comes to my attention while at **[name of registry].** I will do this in accordance with the **[name of registry]'**s privacy policy and applicable laws, including those that require mandatory reporting.

I also agree to never remove any confidential material of any kind from the premises of **[name of registry]** unless authorized as part of my duties, or with the express permission or direction to do so from **[registry director].**

_____ (Print Staff Name)

_____ (Signature of Staff)

_____ (Signature of witness)

Dated this _____ day of _____, 2_____

APPENDIX 4. APPLICATION FORM FOR RELEASE OF DATA

XXXXXXXXXXXXXXX CANCER REGISTRY

APPLICATION FORM FOR RELEASE OF DATA

This form should be completed and accompany each written request for release of cancer registry data. All documents must be submitted to the Registry Director responsible for official approval.

ORGANIZATION OR INDIVIDUAL REQUESTING THE DATA

Name:_____

_

Address:_____

Telephone:_____

Email:_____ Date data are needed:_____

PURPOSE OF DATA

☐ Research ☐ Clinical ☐ Planning ☐ Education ☐ Media

☐ Others, specify _____

Please provide supporting documents for the requested data, including a copy of the research/project proposal and approval from the Research & Ethics Committee if the data is going to be used for research.

DATA ITEMS REQUESTED

Please list the specific data items and year(s) of data requested - use a separate document if necessary.

_____

_____

_____

_____

_____

ASSURANCES

The requested data will not be used for purposes other than those stated above and will not be released to other person(s) or purposes not stated herein. All released data should be destroyed at the end of the activity/project.

All data from the Seychelles National Cancer Registry (SNCR) cited in any presentation or publication should be acknowledged, and the SNCR should receive a copy of the publication upon release.

Name and Signature of Person Requesting Data: _____

Date of Request _____

FOR OFFICE USE ONLY

Approved_____ Date:_____

Processed by:_____ Date:_____

(Name, Signature & Designation)