Cancer incidence, survival and prevalence statistics

Methodology report

Northern Ireland Cancer Registry, 2023

An official statistics publication

ABOUT THIS REPORT

Contents

This report includes information on the methodology used by the Northern Ireland Cancer Registry (NICR) in the production of statistics on cancer incidence, survival and prevalence.

Official statistics

Legislation designating NICR as an official producer of statistics came into place on 1st April 2012 [1], with the incidence, prevalence and survival statistics released by the registry designated as Official Statistics. This designation signifies that the production of these statistics comply with the Code of Practice for Official Statistics as set out by the UK Statistics Authority Code of Practice [2].

The framework for the Code of Practice for Statistics is based on trustworthiness (i.e. having confidence in the organisation that produces the statistics), quality (i.e. using data and methods that produce reliable statistics) and value (i.e. the statistics support user's need for information). As a result, official statistics are produced free from political interference and comply fully with the principles and supporting protocols in the code, including the protocol on release practices.

Cancer mortality data

The NI Statistics and Research Agency (NISRA) is the official statistics provider of cancer mortality data in Northern Ireland. However, for completeness, data on cancer mortality is usually provided by NICR alongside the release of cancer incidence data. This report thus includes information on the methodology used in the production of mortality statistics that are produced and released by NICR. Mortality data is provided courtesy of the General Register Office (NI) via the Department of Health.

Reuse of information

The information in this report is available for reuse free of charge and without the need to contact NICR. However, we request that NICR is acknowledged as the source of any reused information. The following reference is recommended:

Northern Ireland Cancer Registry 2023. Cancer incidence, survival and prevalence statistics: Methodology report. Available at: www.qub.ac.uk/research-centres/nicr

Further information

Further information is available at: www.qub.ac.uk/research-centres/nicr

Phone: +44 (0)28 9097 6028 **e-mail:** nicr@qub.ac.uk

Acknowledgements

The Northern Ireland Cancer Registry (NICR) uses data provided by patients and collected by the health service as part of their care and support.

NICR is funded by the Public Health Agency and is based in Queen's University, Belfast.

CANCER REGISTRATION

The Northern Ireland Cancer Registry (NICR) is part of Queen's University, Belfast and is funded by the Public Health Agency to collate information on all new diagnoses of cancer in Northern Ireland (NI). It was first established in 1994 and uses an automated computer system with multiple information sources from across the National Health Service (NHS) in NI to provide detailed information on cancer incidence in NI from 1993 onwards.

REGISTRY OPERATIONS

The NICR acquires notifications of possible cancer and pre-malignant conditions within the NI population from three main sources:

- Pathology reports from the four pathology laboratories in NI (Belfast, Altnagelvin, Antrim and Craigavon);
- Hospital admissions and discharges recorded in the Patient Administration System (PAS) and supplied by the five Health and Social Care Trusts (HSCT); and
- Death registrations from the General Registrar Office (GRO), which are received via the Department of Health (DoH).

These data sources are loaded and combined electronically, with automatic routines applied which cross check key details and resolve multiple notifications. However, considerable manual work is also required to ensure that key data items (e.g. date of diagnosis, cancer type) are coded to international cancer registration standards and that the final data is as complete and as accurate as possible.

As part of this process, a major focus of the registry's operation is on the verification of any registration which comes from a single hospital admission, a single pathology report or a single death certificate. For these registrations trained Cancer Intelligence Officers (CIOs) examine general practitioners' (GPs) notes for patients who have died from cancer, hospital records for cases identified without histopathology or cytology confirmation and pathology reports where there is conflicting information or other possible errors.

To further check the accuracy of any coding, ensure that no duplicate registrations are present and to separate primary cancers from metastatic disease and recurrences, the CIOs review each cancer registration using a range of health care systems. These include:

- Cancer Patient Pathway System (CaPPS);
- Regional Information System for Oncology & Haematology (RISOH);
- Radiotherapy Information System (ARIA); and
- Theatre Management System (TMS)

As part of this review these data sources are also used to add supplementary information to each cancer registration, including stage of disease and cancer specific information (e.g. ER/HER status for breast cancer, Gleason score for prostate cancer) as the majority of this information is not available as part of electronic downloads and may require reading of pathology reports or hospital notes by the CIOs.

DATE OF DIAGNOSIS

One of the primary data items recorded as part of the cancer registration process is the date of cancer diagnosis. NICR base the collection of this data item on recommendations from the European Network of Cancer Registries [3], which states that where possible the date of diagnosis should be the date of first histological or cytological confirmation of the malignancy. Given that this process can involve various stages, the date is chosen according to the following priority:

1. Date when the biopsy was taken;
2. Date of receipt of the sample by the pathologist;
3. Date of the pathology report.

In the scenario where the cancer is not diagnosed pathologically then the date of admission to hospital as a result of this malignancy is used as the date of diagnosis. If no information is available other than the fact that the patient has died as a result of cancer then date of death is used as the date of diagnosis, and the registration is flagged as being death certificate only (DCO).

CANCER CLASSIFICATION AND CODING

Cancer is coded to the tenth revision of the International Classification of Diseases (ICD10) [4]. The ICD10 codes used to classify cancer are C00-C97. Non-melanoma skin cancer (ICD10 code C44) is frequently excluded from the overall cancer count as it is easily treated and rarely fatal.

CANCER TYPE

The ICD10 codes used to classify each type of cancer are listed in table 1.

Table 1: Classification of cancer type based upon ICD10 code

Cancer type	ICD10 code	Cancer type	ICD10 code
Bladder	C67	Lymphoma	C81-C86
Brain (inc. CNS)	C70-C72, C75.1-C75.3	- Hodgkin's lymphoma	- C81
Cervical	C53	- Non-Hodgkin's lymphoma	- C82-C86
Colorectal	C18-C20	Malignant melanoma	C43
- Colon	- C18	Multiple myeloma	C90
- Rectum	- C19-C20	Non-melanoma skin	C44
Breast	C50	Oesophageal	C15
Gallbladder (inc. biliary)	C23-C24	Ovarian (inc. fallopian tube)	C56-C57.4
Head and Neck	C00-C14, C30-C32	Pancreatic	C25
- Oral	- C00-C14	Prostate	C61
- Laryngeal	- C32	Stomach	C16
Kidney	C64	Testicular	C62
Leukaemia	C91-C95	Thyroid	C73
Liver	C22	Unknown primary	C77-C80
Lung (inc. trachea)	C33-C34	Uterine	C54-C55

CNS: Central Nervous System

Cancer subtype

For head and neck cancer a further breakdown into specific cancer types is also provided. This breakdown in also derived using the ICD10 classification

Table 2: Classification of head and neck cancer type based upon ICD10 code

Cancer type	ICD10 code	Cancer type	ICD10 code
Lip cancer	C00	Nasopharyngeal cancer	C11
Tongue cancer	C01-C02	Hypopharyngeal cancer	C12-C13
Mouth cancer	C03-C06	Other mouth/pharyngeal cancer	C14
Cancer of the salivary glands	C07-C08	Laryngeal cancer	C32
Oropharyngeal cancer	C09-C10	Cancer of the nasal cavity or sinuses	C30-C31

CANCER MORPHOLOGY

In addition to the ICD10 classification, the more detailed International Classification of Diseases for Oncology, version 3 (ICD-O3) [5] is used in the coding of tumour site (topography), histology (morphology) and behaviour. This information is usually obtained from the pathology report and allows cancer registries to

distinguish between the likes of adenocarcinoma and squamous cell carcinoma of the lung, which would both use the same code in the ICD10 classification.

Neuroendocrine cancers

Neuroendocrine tumours and neoplasms develop in endocrine and/or nerve cells and can thus occur in most of the bodies organs. Neuroendocrine cancer has malignant characteristics, meaning that it can spread beyond the original location that it developed. To identify neuroendocrine cancer, we use the ICD-O3 and ICD10 combination of codes specified by Genus et al [6] in the first population based study of these cancers conducted in the UK. The codes used are tabulated below.

Table 3: Classification of neuroendocrine cancers based upon ICD10 and ICD-O3 codes

ICD10 code	ICD-O3 morphology code
C34, C78	8150, 8151, 8152, 8153, 8154, 8155, 8156, 8157, 8158, 8240, 8241, 8242, 8243, 8244, 8245, 8246, 8247, 8249, 9091
C00-C32, C37-C77, C79-C80	8013, 8041, 8042, 8043, 8044, 8045, 8150, 8151, 8152, 8153, 8154, 8155, 8156, 8157, 8158, 8240, 8241, 8242, 8243, 8244, 8245, 8246, 8247, 8249, 9091

Sarcomas

Sarcomas are cancers that develop either in the connective and supporting tissues of the body (such as muscles and nerves) or in the bone. Given that sarcomas can occur in any part of the body, the ICD-O3 classification is used to identify these types of cancer. The codes used are based upon those identified by the National Disease Registration Service as part of their "Get Data Out" programme [7]. The codes used are tabulated below.

Table 4: Classification of sarcoma based upon ICD-O3 code

Sarcoma type	Behaviour code	Topography code	Morphology code
Bone	3	C00-C80	9195, 9221, 9230, 9364
	3	C00-C80, ex. C51-C58	9260, 9365, 9370, 9371, 9372
	3	C40	8902, 8910
	3	C40-C41	8003, 8800, 8801, 8802, 8803, 8804, 8805, 8806, 8810, 8811, 8812, 8813, 8815, 8825, 8830, 8832, 8833, 8840, 8850, 8851, 8852, 8853, 8854, 8855, 8857, 8858, 8860, 8890, 8891, 8894, 8896, 8897, 8901, 8921, 8982, 8990, 8991, 9044, 9120, 9130, 9133, 9150, 9170, 9180, 9181, 9182, 9183, 9184, 9185, 9186, 9187, 9192, 9193, 9194, 9220, 9231, 9240, 9242, 9243, 9250, 9251, 9252, 9540, 9560, 9561, 9571, 9580
	6	C00-C80	9180, 9181, 9185, 9194, 9220, 9221, 9230, 9242, 9243
	6	C00-C80, ex. C40-C41	9364
	6	C00-C80, ex. C51-C58	8812, 9260, 9365, 9370, 9371, 9372
	6	C40-C41	9182, 9184, 9186, 9187, 9192, 9193, 9231, 9240, 9250, 9364
	9	C40-C41	8800, 8890, 9120
Soft tissue	3	C00-C80	8821, 8822, 8933, 9040, 9041, 9042, 9043
	3	C00-C80, ex. C40-C41	8711, 8800, 8801, 8802, 8803, 8804, 8805, 8806, 8810, 8811, 8813, 8814, 8815, 8824, 8825, 8830, 8832, 8840, 8841, 8850, 8851, 8852, 8853, 8854, 8855, 8857, 8858, 8860, 8890, 8891, 8894, 8895, 8896, 8897, 8900, 8901, 8902, 8910, 8912, 8920, 8921, 8930, 8931, 8935, 8940, 8963, 8964, 8982, 8990, 8991, 9044, 9120, 9130, 9133, 9150, 9170, 9180, 9181, 9185, 9194, 9220, 9231, 9240, 9242, 9243, 9251, 9252, 9540, 9560, 9561, 9571, 9580, 9581
	3	C00-C80, ex. C51-C58	8812

	3	C00-C80, ex. C52-C58	8833
	3	C49	8003, 8004
	3	C50, C61	9020
	3	C77	9250
	6	C00-C80	8711, 8800, 8801, 8802, 8803, 8804, 8805, 8806, 8810, 8811, 8813, 8814, 8815, 8821, 8822, 8825, 8830, 8832, 8840, 8841, 8850, 8851, 8852, 8853, 8854, 8855, 8857, 8858, 8860, 8890, 8891, 8894, 8896, 8897, 8900, 8901, 8902, 8910, 8912, 8921, 8930, 8931, 8933, 8935, 8940, 8963, 8964, 8982, 8990, 8991, 9040, 9041, 9042, 9043, 9044, 9120, 9130, 9133, 9150, 9170, 9251, 9252, 9540, 9560, 9561, 9571, 9580, 9581
	6	C00-C80, ex. C40	8895
	6	C00-C80, ex. C40-C41	9231, 9240
	6	C00-C80, ex. C51-C58	8920
	6	C00-C80, ex. C52-C58	8833
	6	C40-C41	9020, 9183
	6	C50, C61	9020
	9	C00-C80, ex. C40-C41	8800, 8890, 8920, 9120
Gastro-intestinal stromal	3,6	C00-C80	8936

Other classifications of sarcoma also exist which may result in slightly different number of cases. Variations are usually a result of how cancers with a mixed sarcoma/epithelial nature are treated (e.g. Mullerian mixed tumour, ICD-O3 morphology 8950).

International classification of childhood cancer

The International Classification of Childhood Cancer (Third edition) [8] was modified and expanded in the UK childhood cancer report [9] to accommodate morphology codes that have been introduced in the first and second revisions of ICD-O3; and to cover the most common sites for tumours in the teenage and young adult age range thereby allowing it's use for a broader age group. The codes used are detailed in table 5.

Table 5: International classification of childhood cancer based upon ICD-O3 code

Diagnostic group	Behaviour code	Topography code	Morphology code
01. Leukaemia, myeloproliferative disease, and myelodysplastic disease			
01a. Lymphoid leukaemia	3,6,9	C00-C80	9820, 9823, 9826, 9827, 9831-9836, 9940, 9948
	3,6,9	C42.0-C42.4, C80.9	9811-9819, 9837
01b. Acute myeloid leukaemia	3,6,9	C00-C80	9840, 9861, 9865-9874, 9877-9879, 9891, 9895-9898, 9910-9912, 9920, 9931
01c. Chronic myeloproliferative disease	3,6,9	C00-C80	9863, 9875, 9876, 9960-9964
01d. Myelodysplastic syndrome and other myeloproliferative disease	3,6,9	C00-C80	9945, 9946, 9975, 9980-9989, 9991-9993
01e. Unspecified and other specified leukaemia	3,6,9	C00-C80	9800, 9801, 9805-9809, 9860, 9930, 9965-9968
02. Lymphoma and reticuloendothelial neoplasm			
02a. Hodgkin's lymphoma	3,6,9	C00-C80	9650-9656, 9659, 9661-9665, 9667
02b. Non-Hodgkin's lymphoma (excluding Burkitt lymphoma)	3,6,9	C00-C80	9591, 9597, 9670, 9671, 9673, 9675, 9678-9680, 9684, 9688-9691, 9695, 9698-9702, 9705, 9708, 9709, 9712, 9714-9719, 9724-9729, 9731-9735, 9737, 9738, 9740-9742, 9750, 9751, 9754-9759, 9760-9762, 9764-9769, 9970, 9971

	3,6,9	C00-C41, C42.2, C44-C77	9811-9819, 9837
02c. Burkitt lymphoma	3,6,9	C00-C80	9687
02d. Miscellaneous lymphoreticular neoplasm	3,6,9	C00-C80	9740-9742, 9749, 9750, 9755-9759
02e. Unspecified lymphoma	3,6,9	C00-C80	9590, 9596
03. Central nervous system and miscellaneous intracranial and intraspinal neoplasm			
03a. Ependymoma and choroid plexus tumour	3,6,9	C00-C80	9383, 9390-9394, 9396
03b. Astrocytoma	3,6,9	C00-C80	9384, 9400-9411, 9420-9425, 9440-9442, 9445
	3,6,9	C72.3	9380
03c. Intracranial and intraspinal embryonal tumour	3,6,9	C00-C80	9470-9478, 9480, 9508
	3,6,9	C70-C72	9501-9504
03d. Other glioma	3,6,9	C00-C80	9381, 9382, 9385, 9430-9432, 9444, 9450, 9451, 9460
	3,6,9	C70, C71, C72.0-C72.2, C72.4-C72.9, C75.1, C75.3	9380
03e. Other specified intracranial and intraspinal neoplasm	3,6,9	C00-C80	8270-8281, 8300, 9350-9352, 9360-9362, 9395, 9412, 9413, 9492, 9493, 9505-9507, 9509, 9530-9539, 9582
03f. Unspecified intracranial and intraspinal neoplasm	3,6,9	C70-C72, C75.1-C75.3	8000-8005
04. Neuroblastoma and other peripheral nervous cell tumours			
04a. Neuroblastoma and ganglioneuroblastoma	3,6,9	C00-C80	9490, 9500
04b. Other peripheral nervous cell tumour	3,6,9	C00-C80	8680-8683, 8690-8693, 8700, 9520-9523
	3,6,9	C00-C69, C73-C76, C80	9501-9504
05. Retinoblastoma			
05. Retinoblastoma	3,6,9	C00-C80	9510-9514
06. Renal tumour			
06a. Nephroblastoma and other non-epithelial renal tumour	3,6,9	C00-C80	8959, 8960, 8964-8967
	3,6,9	C64.9	8963, 9364
06b. Renal carcinoma	3,6,9	C00-C80	8311, 8312, 8316-8319, 8361
	3,6,9	C64.9	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8143, 8155, 8190-8201, 8210, 8211, 8221-8231, 8240, 8241, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8401, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8576
06c. Unspecified malignant renal tumour	3,6,9	C64.9	8000-8005
07. Hepatic tumour			
07a. Hepatoblastoma	3,6,9	C00-C80	8970
07b. Hepatic carcinoma	3,6,9	C00-C80	8160-8180
	3,6,9	C22.0, C22.1	8010-8041, 8050-8075, 8082, 8120-8122, 8140, 8141, 8143, 8155, 8190-8201, 8210, 8211, 8230, 8231, 8240, 8241, 8244-8246, 8260-8264, 8310, 8320, 8323, 8401, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8576
07c. Unspecified malignant hepatic tumour	3,6,9	C22.0, C22.1	8000-8005

08. Malignant bone tumour			
08a. Osteosarcoma	3,6,9	C40, C41, C76, C80	9180-9187, 9191-9195, 9200
08b. Chondrosarcoma	3,6,9	C00-C80	9221, 9230, 9241-9243
	3,6,9	C40, C41, C76, C80	9210, 9220, 9240
08c. Ewing tumour and related sarcoma of bone	3,6,9	C40, C41	9363-9365
	3,6,9	C40, C41, C76, C80	9260
08d. Other specified malignant bone tumour	3,6,9	C00-C80	8812, 9250, 9261, 9262, 9270-9275, 9280-9282, 9290, 9300-9302, 9310-9312, 9320-9322, 9330, 9340-9342, 9370-9372
	3,6,9	C40, C41	8810, 8811, 8814, 8823, 8830
08e. Unspecified malignant bone tumour	3,6,9	C40, C41	8000-8005, 8800, 8801, 8803-8805
09. Soft tissue and other extraosseous sarcoma			
09a. Rhabdomyosarcoma	3,6,9	C00-C80	8900-8905, 8910, 8912, 8920, 8991
09b. Fibrosarcoma, peripheral nerve sheath tumour, and other fibrous neoplasm	3,6,9	C00-C80	8820, 8822, 8824-8827, 9150, 9160, 9491, 9540-9571, 9580
	3,6,9	C00-C39, C44-C76, C80	8810, 8811, 8813-8815, 8821, 8823, 8834, 8835
09c. Kaposi sarcoma	3,6,9	C00-C80	9140
09d. Other specified soft tissue sarcoma	3,6,9	C00-C80	8587, 8710-8714, 8806, 8831-8833, 8836, 8840-8842, 8850-8858, 8860-8862, 8870, 8880, 8881, 8890-8898, 8921, 8990, 9040-9044, 9120-9125, 9130-9133, 9135-9137, 9141, 9142, 9161, 9170-9175, 9231, 9251, 9252, 9373, 9581
	3,6,9	C00-C39, C44-C76, C80	8830
	3,6,9	C00-C63, C65-C69, C73-C76, C80	8963
	3,6,9	C49	9180, 9210, 9220, 9240
	3,6,9	C00-C39, C47-C75	9260
	3,6,9	C00-C39, C47-C63, C65-C69, C73-C76, C80	9364
	3,6,9	C00-C39, C47-C63, C65-C76, C80	9365
09e. Unspecified soft tissue sarcoma	3,6,9	C00-C39, C44-C76, C80	8800-8805
10. Germ cell tumour, trophoblastic tumour, and neoplasm of gonad			
10a. Intracranial and intraspinal germ cell tumour	3,6,9	C70-C72, C75.1-C75.3	9060-9065, 9070-9072, 9080-9085, 9100, 9101
10b. Malignant extracranial and extragonadal germ cell tumour	3,6,9	C00-C55, C57-C61, C63-C69, C73-C75.0, C75.4-C76, C80	9060-9065, 9070-9072, 9080-9086, 9100-9105
10c. Malignant gonadal germ cell tumour	3,6,9	C56, C62	9060-9065, 9070-9073, 9080-9085, 9090-9093, 9100, 9101
10d. Gonadal carcinoma	3,6,9	C00-C80	8441-8444, 8450, 8451, 8460-8473
	3,6,9	C56, C62	8010-8041, 8044, 8050-8075, 8082, 8120-8123, 8130-8141, 8143, 8190-8201, 8210, 8211, 8221-8241, 8244-8246, 8249, 8260-8263, 8290, 8310, 8313, 8320, 8323, 8380-8384, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8573, 9000, 9014, 9015
10e. Other and unspecified malignant	3,6,9	C00-C80	8590-8671

gonadal tumour	3,6,9	C56, C62	8000-8005
11. Other malignant epithelial neoplasm malignant melanoma			
11a. Adrenocortical carcinoma	3,6,9	C00-C80	8370-8375
	3,6,9	C74	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11b. Thyroid carcinoma	3,6,9	C73	8010-8158, 8190-8231, 8240-8369, 8376-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11c. Nasopharyngeal carcinoma	3,6,9	C11	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11d. Malignant melanoma	3,6,9	C00-C80	8720-8780, 8790
11e. Skin carcinoma	3,6,9	C44	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11f. Colon carcinoma (excluding appendix)	3,6,9	C18, C19, C20 ex. C18.1	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11g. Breast carcinoma	3,6,9	C50	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11h. Cervical carcinoma	3,6,9	C53	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11i. Bladder carcinoma	3,6,9	C67	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11j. Salivary gland carcinoma	3,6,9	C07, C08	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11k. Other head and neck carcinoma	3,6,9	C00-C06, C09, C10, C12-C14, C30-C32, C76.0	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11l. Stomach and upper gastrointestinal carcinoma	3,6,9	C15-C17, C23, C24, C26	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11m. Pancreatic carcinoma	3,6,9	C25	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11n. Lung carcinoma	3,6,9	C33, C34	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11o. Carcinoma of other specified sites	3,6,9	C18.1, C37-C39, C48, C51, C52, C54, C55, C57, C58, C60, C61, C63, C65, C66, C68, C69, C75.0, C75.4-C75.9	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
11p. Carcinoma of unspecified sites	3,6,9	C70-C72, C76.1-C80	8010-8158, 8190-8231, 8240-8369, 8376-8440, 8452-8454, 8480-8589, 8941, 8983, 9000, 9010-9016, 9020, 9030
12. Other and unspecified malignant neoplasms			
12a. Other specified malignant tumour	3,6,9	C00-C80	8930-8936, 8940, 8950, 8951, 8971-8982, 9050-9055, 9110
	3,6,9	C00-C39, C47-C75	9363
12b. Other unspecified malignant tumour	3,6,9	C00-C21, C23-C39, C42-C55, C57-C61, C63, C65-C69, C73-C75.0, C75.4-C80	8000-8005

TUMOUR CHARACTERISTICS

CANCER STAGE

Staging is carried out using a number of laboratory and clinical tests at diagnosis. The staging classification used throughout the official statistics is the TNM stage [10,11] that includes information on the extent of the primary tumour (T), the absence or presence of lymph node metastasis (N) and the absence or presence of distant metastasis (M). The classification combines these three elements to produce an overall TNM stage for the tumour, although the manner in which the overall TNM stage is derived depends upon the cancer site. Staging is carried out on most cancer sites, although for some cancer sites, specifically non-melanoma skin cancer, brain cancer and leukaemia there is no TNM classification.

For analysis purposes the overall TNM stage for each cancer is coded to four groups, ranging from early tumours (Stage I) to advanced tumours that have distant metastasis (Stage IV).

The TNM classification has changed over time. For most cancer types version 7 is used prior to 2018, with version 8 used from 2018 onwards. The sole exception is for head and neck cancer, which uses version 7 prior to 2019, with version 8 used from 2019 onwards.

ADDITIONAL INFORMATION

MORTALITY DATA

Mortality data is supplied by the General Registrar Office of Northern Ireland (GRONI) via the Department of Health (DoH). It is supplied on a quarterly basis and is used both to determine annual cancer mortality figures, but also to provide follow up information on cancer patients in survival and prevalence calculations. For the purpose of the production of annual mortality data, cancer type is coded using the same ICD10 codes listed in table 1, while the annual count of number of cancer deaths is based upon the year that death occurs. Figures may thus differ slightly from those produced by the NI Statistics and Research Agency who use the year that deaths are registered.

POPULATION DATA

Population data for each calendar year and for a range of geographic areas is required in the calculation of incidence and mortality rates. Mid-year population estimates are produced by the Northern Ireland Statistics and Research Agency (NISRA) on an annual basis and are the source of all population data used throughout the NICR official statistics publications. They are freely available from the NISRA web site [12] and are downloaded from this site by NICR on an annual basis.

Mid-year population estimates refer to the usually resident population in Northern Ireland at 30 June each year, with the usually resident population defined according to United Nations guidelines. The estimates are derived using the components of change method, where starting with the most recent census figure the population estimate is aged by one year, with births added, deaths removed and net migration accounted for.

GEOGRAPHIC AREAS

NICR routinely collects address information, including postcode, allowing geographic areas to be assigned to records of cancer incidence. This is accomplished for each patient through an electronic process that uses the collected postcode along with a lookup file, known as the Central Postcode Directory (CPD) [13], which provides the relationship between each valid postcode in Northern Ireland and a range of higher geographic areas. The key areas derived from the patient's postcode in this manner are Health and Social Care Trusts (HSCT), Local Government Districts (LGD), Parliamentary Constituencies and Super Output Areas (SOA - a small geographic area with a target population of around 2,000 people). Addresses with an unknown, incomplete or invalid postcode cannot be assigned higher geographic areas, however only a small proportion of records for cancers diagnosed fall into this category.

SOCIO-ECONOMIC DEPRIVATION

The 2017 Northern Ireland multiple deprivation measure (NIMDM) [14] assigns a deprivation score to each Super Output Area (SOA) in Northern Ireland based upon the economic characteristics of all persons usually resident in that area. For the purposes of NICR official statistics publications the SOAs were ranked according to this score and divided into quintiles, with quintile 1 containing the fifth of the population resident in the most deprived SOAs and quintile 5 containing the fifth of the population resident in the least deprived SOAs. Patients were then assigned a deprivation quintile based upon their SOA of residence which was derived for each patient based upon their postcode of residence.

MEASURES OF CANCER INCIDENCE AND MORTALITY

The statistical methods used by NICR are those widely used by cancer registries throughout the world in describing the burden of cancer within their catchment area. Consequently a significant amount of literature is available on these techniques (e.g. 15,16). This section thus only provides a very general overview of the techniques utilized and is not meant to supplant the numerous (and better) texts on cancer registration techniques and medical statistics.

DESCRIPTIVE STATISTICS

The most common and useful measure of cancer levels in a population is the absolute number of cases (incidence) or deaths (mortality) in a given period of time. It is these very basic figures that allow planning by the health service and are the fundamental building blocks of any other analysis.

Mean (Average) number of cases/deaths

However, the number of cancer cases/deaths within a year compared to the size of the population of Northern Ireland is relatively small. This can result in the number of events being studied fluctuating each year as a result of natural variation, particularly when data are broken down by smaller geographic areas such as Health Trusts or by patient demographics such as age. In order to introduce more stability into any presented statistics we typically observe the population over five or more years and present a mean number of cases/deaths per year, which should be interpreted as a typical value for the annual number of cases/deaths in the patient group being studied.

Proportions and rates

In order to properly investigate the distribution of cancer by demographic and tumour characteristics and to make comparisons between different groups, proportions or rates are presented alongside most counts of events. Proportions are used whenever the total being used is a fixed cancer-based population (e.g. total number of cancer deaths, total number of male lung cancer cases) which is derived from the same source data as the number of events. All proportions are multiplied by 100% to provide a percentage value.

Rates are used whenever the population of interest is the general population, most of which is cancer free. The population denominators in this report come from the mid-year population estimates [12] which match the groups of interest (e.g. males aged 40-45 resident in Mid-Ulster Local Government District in 2020). Rates are typically multiplied by 100,000 to provide a readable statistic and are referred to as a crude incidence/ mortality rate per 100,000 persons.

AGE-STANDARDISED RATES

When comparing one or more patient groups, crude rates are not always the best measure to use as there is a very strong relationship between cancer and age, thus a younger population is more likely to have a lower number of cancers than an older population of the same size. The most useful and easiest to calculate

measure that compensates for differences in the age structures of two populations is a set of age-specific rates, which are calculated in a similar manner to crude rates by dividing the number of events by the population of a specific age group and multiplying by 100,000.

The drawback of the use of age-specific rates is the number of these that must be quoted in order to give a full picture of the cancer/population being studied, particularly since five-year age-groups are the most commonly used age breakdown. In addition the small numbers involved can cause very noticeable fluctuations over time, even when several years' worth of data are used.

A widely used technique, which provides a summary measure that allows for the changing or differing population age-structure, is age-standardisation. This does not completely overcome the difficulty in comparing rates between populations and is thus not a replacement for age-specific rates but does provide statistics that are more manageable and lend themselves to further analysis, particularly comparisons of many sets of incidence/mortality rates such as is required in trend and geographic analysis.

There are two methods of age-standardisation, direct and indirect, used in analysis of cancer incidence and mortality. The former provides an absolute measure (i.e. the number of events per 100,000 persons), while the indirect method provides a value relative to some other measure (i.e. the number of observed events compared to the number of expected events).

Direct standardisation

The result of direct standardisation is known as an age-standardised rate (ASR), which refers to the number of events per 100,000 persons occurring in the population if the population possessed the same age structure as a standard population. The standard population used by NICR in incidence and mortality analysis is the 2013 European standard population which is tabulated below.

Table 6: 2013 European standard population

Age group	Population	Age group	Population	Age group	Population	Age group	Population
0-4	5,000	**25-29**	6,000	**50-54**	7,000	**75-79**	4,000
5-9	5,500	**30-34**	6,500	**55-59**	6,500	**80-84**	2,500
10-14	5,500	**35-39**	7,000	**60-64**	6,000	**85-89**	1,500
15-19	5,500	**40-44**	7,000	**65-69**	5,500	**90+**	1,000
20-24	6,000	**45-49**	7,000	**70-74**	5,000		

The calculation of an age-standardised rate is based upon the age-specific rates. These rates are multiplied by the standard population for that age class (also known as the weight), with the products summed and divided by the total standard population. It is typically multiplied by 100,000. This value is the standard measure used for making comparisons between different populations, however while useful within this context it cannot be interpreted as a measure of the actual number of events within a population due to the removal of the age effect and corresponds to the crude rate in the standard population rather than that being studied.

Indirect standardisation

The indirect method of age-standardisation is a comparison of the observed number of events within a population and the number of events expected in a reference population of the same size. When considering incidence of cancer the expected number is calculated by applying the age specific rates of a reference

population to the observed population (i.e. the population being studied). The result is usually expressed as a percentage by multiplying by 100, with 100% referring to the events in the reference population and the result known as the standardised incidence ratio (SIR). This approach is also valid for mortality data using age-specific mortality rates with the result known as the standardised mortality ratio (SMR).

This measure is frequently used for geographic analysis with the reference population being that for an entire country (i.e. Northern Ireland) and SIRs or SMRs calculated for smaller geographic units (e.g. district councils/deprivation quintiles) giving an indication of how cancer levels in these areas compare to that of the entire country.

Cumulative risk/Odds

Another commonly used measure which is of particular interest to the general public, but is not as useful as age-standardised rates, is the cumulative risk, which gives the risk of an individual developing cancer during a particular age span (i.e. before ages 75 or 85) assuming the absence of other causes of death. Like age-standardised rates it is based upon age-specific rates but is expressed as a percentage rather than a rate. Dividing 1 by the cumulative risk gives the odds that the event will happen before the given age.

CONFIDENCE INTERVALS AND STATISTICAL SIGNIFICANCE

However, one important feature of age-standardised rates is that they are only estimates of the true value, as uncertainty exists due to random fluctuations in the number of events between different populations. In order to quantify this uncertainty any rates are accompanied by 95% confidence intervals to indicate the range within which there is a 95% probability that the true value is likely to fall. The size of the confidence interval depends upon the number of events and the size of the population within which they occur, with rates made up of a small number of observations within a large population being less stable and having large confidence intervals.

Rates for two different time periods or population groups are considered to differ only if the 95% confidence intervals for the two age-standardised rates do not overlap. Alternatively, in the case of ratios, the rates for a population differ from those of the reference population only if the confidence interval does not include 100%. This is known as statistical significance and for significant differences the level of certainty about any difference can be quantified by calculating the p-value. This measure provides the probability that any difference observed between two rates is due to chance. Thus a p-value of 0.001 indicates a 99.9% probability that differences are genuine and not a result of random factors.

CANCER SURVIVAL

Survival refers to the proportion of patients who survive a given amount of time after a diagnosis of cancer. It is one of the best indicators as to the efficiency of diagnostic and treatment methods in a geographic area and is widely used by cancer registries as a broad indicator as to the effectiveness of health services in the treatment of cancer.

It is evaluated using two measures. Observed survival examines the time between diagnosis and death from any cause. It thus represents what cancer patients experience, however, due to the inclusion of non-cancer deaths (e.g. heart disease), it may not reflect how changes in cancer care impact survival from cancer. Thus age-standardised net survival is also examined. This measure provides an estimate of patient survival which has been adjusted to take account of deaths unrelated to cancer. It also assumes a standard age distribution thereby removing the impact of changes in the age distribution of cancer patients on changes in survival over time. While this measure is hypothetical, as it assumes patients can only die from cancer related factors, it is a better indicator of the impact of changes in cancer care on patient survival.

Observed survival

Observed survival is the proportion of patients who survive a specified amount of time after the point that they are diagnosed with cancer. It is typically presented as the probability of a group of patients surviving a specified amount of time regardless of the cause of death. The observed survival of a group of patients is determined by first assessing whether each patient in the group is alive or dead at a certain date. This date is known as the censor date, while the alive or dead status of each patient at this time is known as the vital status. If a patient's vital status is unknown at the censor date (e.g. if they have moved abroad) the censor date is changed to be the date that the patients' vital status was last recorded. Such patients are referred to as being lost to follow up. The number of patients recorded as lost to follow up in Northern Ireland is small due to the use of the passive follow up method of data collection.

If the patient has been identified as having died between the date of diagnosis and the censor date the observed survival time is calculated as the difference, in days, between the exact date of diagnosis and the exact date of death as recorded by the GRO. If the patient is still alive at the censor date the survival time is the difference between the date of diagnosis and the censor date. Such patients are known as "censored" patients. The calculated survival time and vital status are then used in the Kaplan-Meier approach to produce survival estimates at various points after diagnosis, up to a maximum of five years.

While all patients are typically included in incidence calculations, survival calculations are restricted to adults aged 15 to 99 at the time of diagnosis with patients aged 100 and over excluded due to problems with data quality and interpretation of the results. Survival among children aged 0 to 14 is studied and presented separately. In addition death certificate only (DCO) registrations are excluded as these registrations have the same date of diagnosis as their date of death with a resultant zero survival time which would bias survival rates downward. In situations where a patient has multiple tumours of the type being studied only the first tumour diagnosed is included in the survival analysis.

Net survival

Observed survival for cancer patients includes death from causes other than cancer, some of which may be related to cancer or its cause (e.g. other smoking related illnesses) or may even be completely unconnected to the disease (e.g. accidental death). It is thus not the best survival measure for monitoring the effectiveness of treatment of the disease or its impact on society. Instead measures that remove other causes of death from survival figures are preferred. The most commonly used of these types of measures, but not the only one, is net survival, which is the difference between the observed survival of a given group of patients and the expected survival for a group of persons in the general population with the same characteristics (usually sex and age).

The expected survival can be calculated using several different techniques. The method used in this report is the Pohar-Perme method [17] and is calculated using the stns module in the Stata statistical software package [18]. This requires the use of background mortality rates by calendar year, sex and single year of age. These are derived from mortality data provided by GRO, but are smoothed using Poisson regression in order to remove fluctuations caused by the small number of events recorded.

Age-standardised net survival

Survival from cancer is dependent upon age at diagnosis. Thus when comparing survival from different populations the same difficulties that occur when comparing incidence and mortality rates are also apparent, with populations having high percentages of younger people having better survival that those with high percentages of older people. To compensate we thus apply the direct age-standardisation approach that was used with incidence and mortality rates to net survival rates. As before this involves the application of a standard population to age-specific net survival rates. However it is necessary to bear in mind that the population at risk when investigating survival differs from that for incidence in that the latter refers to the entire population while survival relates only to patients with cancer. Thus different standard populations are required, not only from that used for the age-standardisation of incidence data but also for those cancer sites that have significantly different age distributions for patients than usual, such as testicular cancer which is more predominant in young men and prostate cancer which is more common in the very elderly. In this report we use the same standard populations suggested by Corazziari et al [19], but collapsed to four age groups due to the small number of events in the NI population for specific age ranges.

Table 7: Standard cancer populations used in age-standardisation of net survival

Standard 1		Standard 2		Standard 3		Standard 4	
Age group	**Population**	**Age group**	**Population**	**Age group**	**Population**	**Age group**	**Population**
15-54	19,000	**15-44**	28,000	**15-44**	60,000	**15-64**	42,000
55-64	23,000	**45-54**	17,000	**45-54**	10,000	**65-74**	29,000
65-74	29,000	**55-64**	21,000	**55-64**	10,000	**75-84**	23,000
75+	29,000	**65+**	34,000	**65+**	20,000	**85+**	6,000
Cancer sites							
All except for those requiring standards 2-4		Sarcoma, melanoma, cervix, brain, thyroid, bone		Testes, Hodgkin's disease		Prostate	

Confidence intervals

As with the other statistical measures used in this report observed and net survival values are accompanied by 95% confidence intervals.

PREVALENCE MEASURES

Prevalence refers to the number of people living in a population with a diagnosis of cancer. Most cancer registries have difficulty in providing an exact figure for this value for a variety of reasons. In the context of Northern Ireland the problems are twofold:

1. There is no point at which cancer is considered cured. While some people diagnosed with cancer may be cancer free within a few years, others may need treatment for a considerable length of time. Thus in order to develop prevalence figures, either an assumption must be made as to an average "cure" point (sometimes arbitrarily taken as being five or ten years) or all people who have been diagnosed with cancer and are still alive at a certain point must be included.
2. NICR have information on people diagnosed with cancer from 1993 onward. Unfortunately with regard to measuring prevalence, this means that there is no information on members of the population who had a diagnosis of cancer prior to 1993. Thus any complete prevalence figures produced is an undercount of the true value.

Figures for complete prevalence are thus not provided by NICR, however prevalence figures for people diagnosed within a fixed period of time are provided (i.e. those diagnosed within one, five, ten and twenty-five years). These refer to the number of people who are alive at the end of a specified calendar year and have previously been diagnosed with cancer up to one, five, ten and twenty-five years ago respectively.

Cancer prevalence is based upon patients rather than tumours and only the most recent diagnosed tumour of the cancer type under consideration is counted. Thus if a patient has been diagnosed with one colorectal tumour and one breast tumour since 1993 they contribute to both the colorectal cancer and breast cancer prevalence count, however they contribute only once to the all cancers count. Similarly a patient with two breast cancers since 1993 contributes only once to the breast cancer count.

REFERENCES

1. The Official Statistics Order (Northern Ireland) 2012. Available at https://www.legislation.gov.uk/nisr/2012/2/contents/made
2. UK Statistics Authority. Code of Practice for Statistics. Available at https://code.statisticsauthority.gov.uk
3. Tyczynski JE, Demaret E, Parkin DM. Standards and guidelines for cancer registration in Europe. IARC Technical Publication No. 40. IARC, Lyon, 2003
4. World Health Organisation. ICD10 International Classification of Diseases 10th revision. Geneva: WHO; 1997
5. World Health Organisation. International Classification of Diseases for Oncology Third Edition. Geneva: WHO; 2013
6. Genus TSE, Bouvier C, Wong K, et al. Impact of neuroendocrine morphology on cancer outcomes and stage at diagnosis: a UK nationwide cohort study 2013-2015. BJC 2019; 121:966
7. National Disease Registration Service. Get Data Out: Sarcoma. Available at www.cancerdata.nhs.uk/getdataout/sarcoma
8. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. Cancer 2005; 1457-1467
9. Public Health England. Children, teenagers and young adults UK cancer statistics report 2021. Available at: http://www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/cancer_in_children_teenagers_and_young_adults/
10. Sobin LH, Gospodarowicz MK, Wittekind C, eds. TNM classification of malignant tumors (7th edition). West Sussex UK: Wiley-Blackwell 2010
11. Brierley JD, Gospodarowicz MK, Wittekind C, eds. TNM classification of malignant tumors (8th edition). West Sussex UK: Wiley-Blackwell 2017
12. Northern Ireland Statistics and Research Agency. Mid-year population estimates. Available at: www.nisra.gov.uk/statistics/population/mid-year-population-estimates
13. Northern Ireland Statistics and Research Agency. Central Postcode Directory. Available at: www.nisra.gov.uk/support/geography/central-postcode-directory
14. Northern Ireland Statistics and Research Agency. Northern Ireland Multiple Deprivation Measure. Available at: www.nisra.gov.uk/statistics/deprivation
15. Boyle P, Parkin DM. Statistical methods for registries. In: Jensen OM, Parkin DM, MacClennan R, et al. Cancer registration principles and methods. Lyon: IARCPress; 1991
16. Breslow NE, Day NE. Statistical methods in cancer research, volume II: The design and analysis of cohort studies. Lyon: International Agency for Research on Cancer, World Health Organization; 1987
17. Perme MP, Stare J, Esteve J. On estimation in relative survival. Biometrics 2012; 68:113-20
18. Clerc-Urmès I, Grzebyk M, Hédelin, G. Net survival estimation with stns. Stata Journal 2014; 14: 87-102
19. Corazziari I, Quinn M and Capocaccia R. Standard cancer population for estimating age standardising survival ratios. Eur J Cancer 2004; 40: 2307-16