

# Data Quality and Completeness at the Irish National Cancer Registry

## Contents

- Data Quality and Completeness at the Irish National Cancer Registry ..... 1
- 1. Introduction ..... 2
- 2. Material..... 2
- 3. Methods..... 3
  - 3.1 Comparability..... 3
  - 3.2 Completeness..... 3
  - 3.3 Validity ..... 4
- 4. Results..... 5
  - 4.1 Comparability..... 5
  - 4.2 Completeness..... 5
    - Completeness: semi-quantative methods ..... 5
    - Completeness : quantative methods..... 8
  - 4.3 Validity ..... 10
- 4 Discussion..... 12
- Bibliography ..... 14

## 1. Introduction

The Irish National Cancer Registry (NCR) is a statutory body that began full registration of all cancers in the Republic of Ireland in January 1994. The reporting of cancer is not mandatory in Ireland. However the NCR makes considerable effort to ensure that there is accurate and complete recording of all cases diagnosed, with active ascertainment and follow-up of cases. The NCR has also recorded all deaths due to cancer since the beginning of 1994 and deaths, from whatever cause, of all patients who have been registered as having cancer. In this study, we aim to give a broad review of data quality at the NCR by examining the comparability, completeness and validity of the data using some of the techniques outlined in papers by Parkin and Bray (1; 2) .

## 2. Material

The registry obtains data on patients with cancer from a variety of sources, primarily via qualified tumour registration officers (TRO) who are employed by the registry and based in hospitals around the country. The main source of notification of each new case is a pathology report.

The second main source of notification is the hospital inpatient enquiry (HIPE) system. Each public hospital produces an annual HIPE list which identifies all cases of cancer discharged from the hospital during that year. A listing of all cancer patients on the HIPE system is provided to the registry on a regular basis. This allows the registry to identify cases that have not had a histological verification of the diagnosis, or for which the registry failed to identify a pathology report.

The third main source comprises death certificates supplied to the NCR by the Central Statistics Office (CSO). These are provided after the CSO has published their quarterly report so a patient will have been deceased at least 6 – 12 months before the death is registered by the NCR. The cancer death certificates are matched against the NCR system and unmatched cases are followed up; if a patient dies in hospital, the death certificate information is sent for checking and follow up to the TRO responsible for registering cases in that hospital; if a patient does not die in a hospital, the physician certifying the death is contacted and can either give the information necessary for a full registration or identify a hospital in which the patient was treated for the cancer. In the latter case, the information is then sent to the relevant tumour registration officer (TRO) at that hospital and the medical records retrieved. If, after follow up, no patient or cancer can be identified which corresponds to the death certificate, the case is registered as a “death certificate only” (DCO) case.

In addition, notifications, updates or treatment of patients normally resident within this Registry area are sometimes received from registries in the UK.

All registrations are checked against the existing Registry database for duplication at the time of entry. Records can usually be matched on the basis of full name, address, date of birth and medical record number. If no previous entry exists for the cancer, it is registered, and the record added to the database. The medical records are requested and retrieved 6–12 months

later to complete cases and capture relevant treatment information. Validation checks are performed at the point of entry. Internal verifications are carried out on a monthly basis. In the preparation of the report data set, the IARC-CHECK program (3) is implemented. This performs validity checks both on individual data items and on consistency between items.

### **3. Methods**

#### **3.1 Comparability**

It is important that information collected by the registry is coded and classified according to international guidelines so that Irish cancer statistics can be compared to other countries worldwide. Issues considered in this area include international classification of diseases (ICD) codes, definitions of incidence and how multiple tumours are handled.

#### **3.2 Completeness**

Completeness is a measure of the extent to which the NCR is capturing diagnosed cancers in Ireland. There are various ways to examine this and the methods fall broadly into two main categories, depending on whether they can be considered semi-quantitative or quantitative. The semi-quantitative methods presented here include: the stability of incidence over time (1994-2008), age-specific incidence rates of childhood cancer (1994-2008), comparisons of the mortality to incidence ratios (2003-2007) between Ireland and Finland and the number of sources per case (2003-2007).

When examining childhood cancer rates, we calculated three-year moving averages which indicate the temporal trend since the annual case numbers are small. A three-year moving average of age specific incidence rates was computed separately for males and females. The moving average was constructed by centring on the mid-year.

The quantitative methods used include independent case ascertainment, the Lincoln-Petersen estimator (3; 4) and the flow method (5).

Independent case-ascertainment is an important quantitative method. As the Registry tries to use all available sources of data, there are few independent registers of cancer cases in Ireland. The national breast screening programme, BreastCheck (6) is the only significant independent national source of which we are aware which is not used for case ascertainment by the Registry. We examined whether cases reported in the BreastCheck programme were already registered by the Registry in the period 2000-2009.

Various other quantitative methods to estimate completeness of case ascertainment have been developed, some of which are based on the number of cases ascertained first from death certificates (DCN cases). These approaches can give an estimate of the number of cases not registered while the individual was alive (3; 5).

We know the mortality to incidence (M:I) ratio (i.e. the ratio of the number of deaths to the number of cases) in registered patients and also the number of DCN cases. The main assumption of the Lincoln-Peterson estimator is that the M:I ratio in

registered patients is the same as the M:I ratio in un-registered patients. This allows us to estimate the number of people who are not registered and have not died of their cancer, as the product of the number of DCNs and the M:I ratio in the registered patients minus the number of DCNs. However, if the M:I ratio in the unregistered population is higher than the registered one, the completeness is bounded as follows:

$$\hat{C}_{LP} \geq \frac{I_r - DCN}{I_r + \frac{DCN}{(M:I)_r}}$$

where  $I_r$  and  $(M:I)_r$  are the incidence and the mortality to incidence ratio respectively for those registered while alive. Thus the estimate calculated using the LP estimator is a lower bound for the completeness of case ascertainment. Clearly the numbers of people diagnosed in a particular year who die of cancer depends on the time since diagnosis, thus this estimate depends on when the figures above are counted. The LP estimate of completeness is made at a fixed point in time, the date of extraction of the data from the Registry database, in this instance 31/12/2010. Thus any cases diagnosed or registered, or deaths observed, after this date were not considered.

The flow method estimates the proportion of people with a cancer diagnosis who are unregistered and a) still alive or b) have died from non-cancer causes. It calculates these estimates using time to event methodology. Software is provided by the authors (5). Completeness can be estimated at any time from the year of diagnosis. We present results for five-year completeness. In addition, one and three year completeness are also shown in

Table 2. For both the LP estimate and the flow method, we calculate completeness for the sample year of 2005. Cases diagnosed in this sample year were extracted at the end of 2010, for all invasive cancers combined (excluding non melanoma skin cancers) and separately for the four most common cancers; lung, breast, colorectal and prostate. We excluded registrations of non-melanoma skin cancer from consideration because these cancers are not registered by most registries.

### 3.3 Validity

Correct information regarding a case is more likely if it has been histologically verified, thus the percentage of cases with histological verification (for cases diagnosed in 2003 - 2007) is presented. If the only information about a case is via a death certificate, then due to the lack of follow-up information we cannot confirm that the case is indeed cancer. Thus a low percentage of DCOs is an important marker for checking the validity of the data. We calculate the percentage of DCO cases for the period 2003-2007. The proportion of registered cases with unknown values for selected data items may be an indicator of data quality. Unknown values may result from problems with data collection, the way code values are defined, or may be due to ambiguity in the medical records which are available. We check the proportion of cases registered with full name and address, date of birth, date of death if relevant, date of incidence and method of diagnosis.

In addition, we check the proportion of cases with primary site unknown (PSU) and the proportion with stage unknown both for the period 2003-2007.

## 4. Results

### 4.1 Comparability

The definition and date of incidence and how multiple primaries are recorded are outlined below.

All cases were registered using the ICD-O2 classification system from 1994-2004, and since 2005 the ICD-O3 system has been used. In the main, data is reported using ICD10 codes. The rules for assigning incidence dates closely follow the European Network of Cancer Registries recommendations which are based on a hierarchy of possible sources. The date of incidence is recorded as one of the following with decreasing order of priority:

1. Date of first histological or cytological diagnosis of the malignancy.
  - a. Date when specimen taken
  - b. Date when specimen examined
  - c. Date of receipt at the laboratory
2. Date of first treatment (excluding “seen but not treated”).
3. Date of admission to hospital because of malignancy.
4. For patients not admitted to hospital, date of first outpatient consultation with the malignancy.
5. Any other date of diagnosis (clinical etc.).
6. Date of death.

However, if the date of first definitive treatment precedes all of these, it is taken as the date of incidence. Only new primary tumours are registered following guidelines developed jointly by the International Association of Cancer Registries (IACR) (7) and International Agency for Research on Cancer (IARC) (8). If a new primary tumour is of the same morphological grouping and site as a primary already registered for the patient, it is flagged as a multiple cancer and discounted from published statistics.

The TNM classification of malignant tumours, fifth edition (9) is used to stage all malignant cancers at sites that have a TNM staging scheme. The system is used to describe the clinical and pathological (usually following surgery) anatomic extent of a particular malignant tumour. After TNM categories are assigned to a tumour they are grouped into stages, 1-4. Once assigned, the registered stage of the tumour remains unchanged.

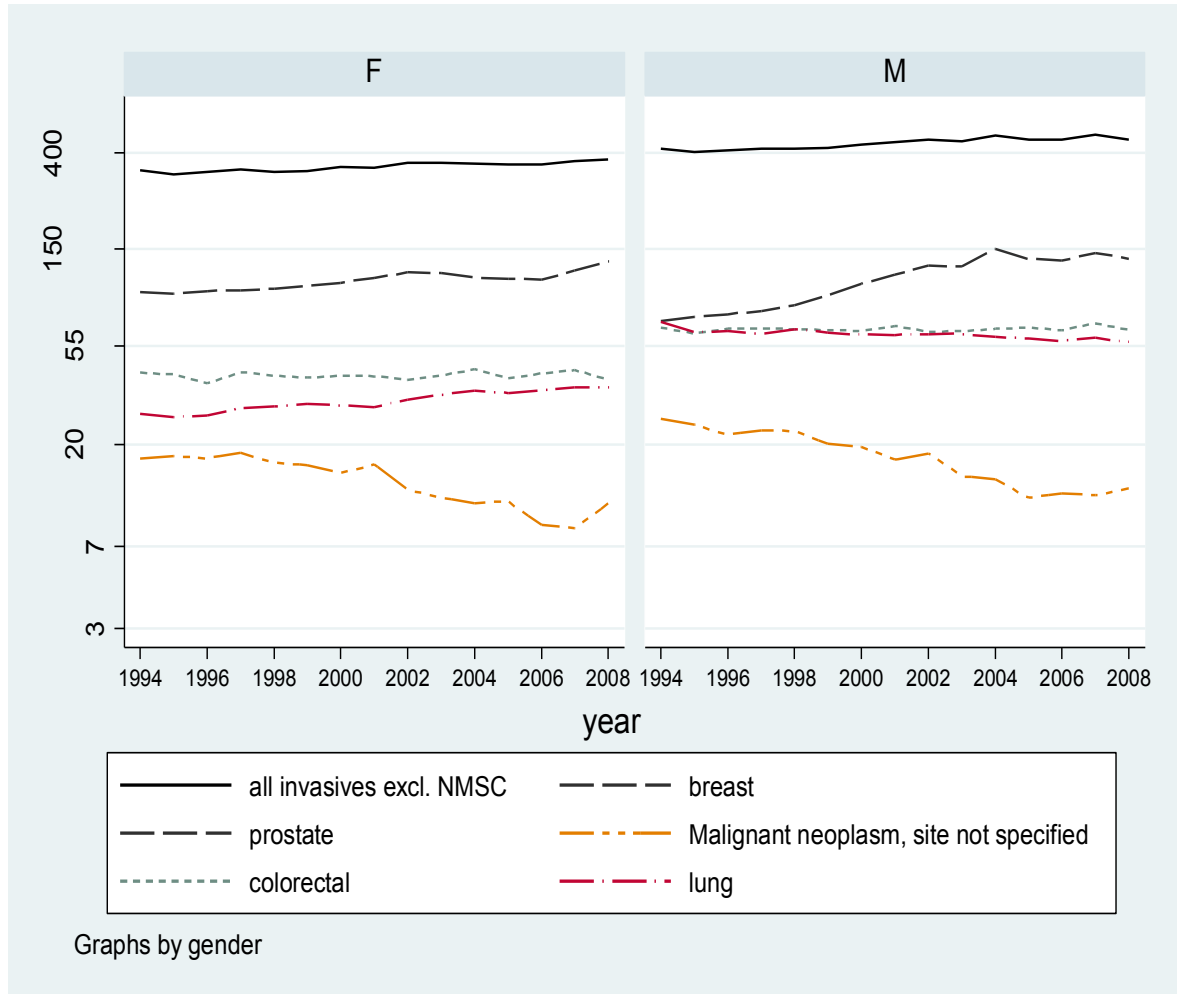
### 4.2 Completeness

#### Completeness: semi-quantative methods

Historical incidence trends for some selected cancers are displayed, by sex in Figure 1. European age standardised incidence rates are presented on a log scale. The overall rates (all invasive cancers excluding non-melanoma skin cancer (NMSC)) for both males and females have remained fairly stable over time. Female breast cancer rates have risen in the last decade, and

this is associated with the national screening programme introduced in this period. For men there was a similar though more dramatic rise in prostate cancer incidence which is linked to an increase in PSA testing. Colorectal rates remained stable for both sexes. There is a decrease in the rates for “primary site unknown”, indicating that more accurate diagnoses/records have been available in more recent years.

Figure 1: Incidence rates (European age-standardised), 1994-2008



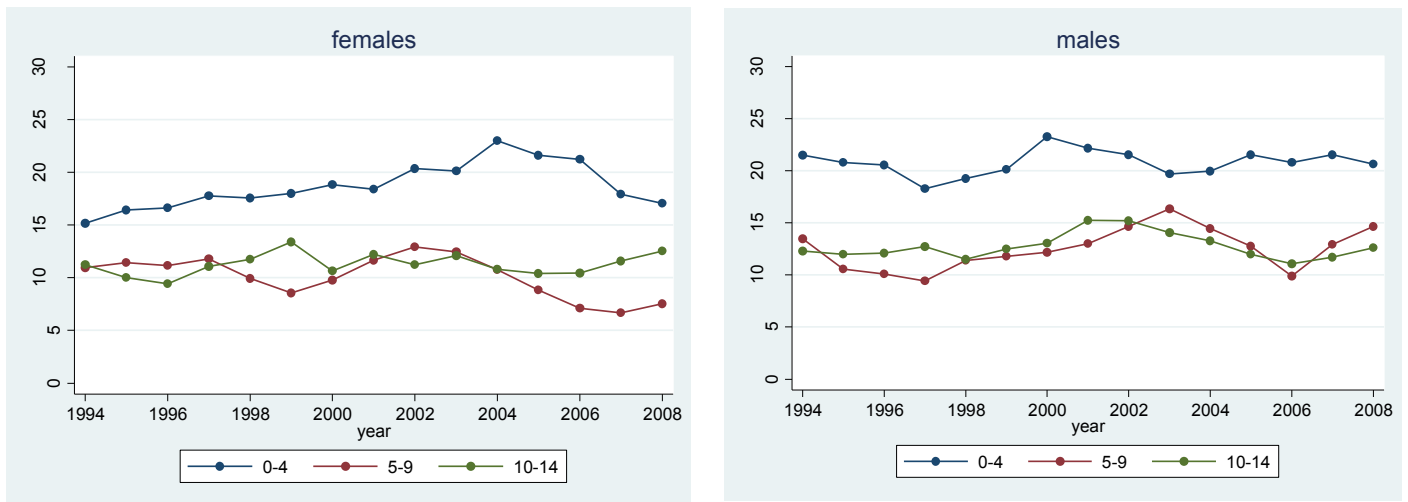
Age specific rates, 0-4, 5-9 and 10-14 years for both childhood cancers for genders in the period 1994-2008 are within an expected range of values taken from table 5.7, chapter 5, Cancer in Five Continents vol. VIII, (10).

Table 1: Age specific incidence rates per 100,000 for childhood cancer by gender, Ireland 1994-2008

Age	girls	reference	boys	reference
0-4	17.7	9.7, 21.4	19.5	12.3, 24.7
5-9	9.2	6.9, 12.0	11.2	8.5, 15.6
10-14	9.9	6.8, 13.6	11.3	8.5, 15.0

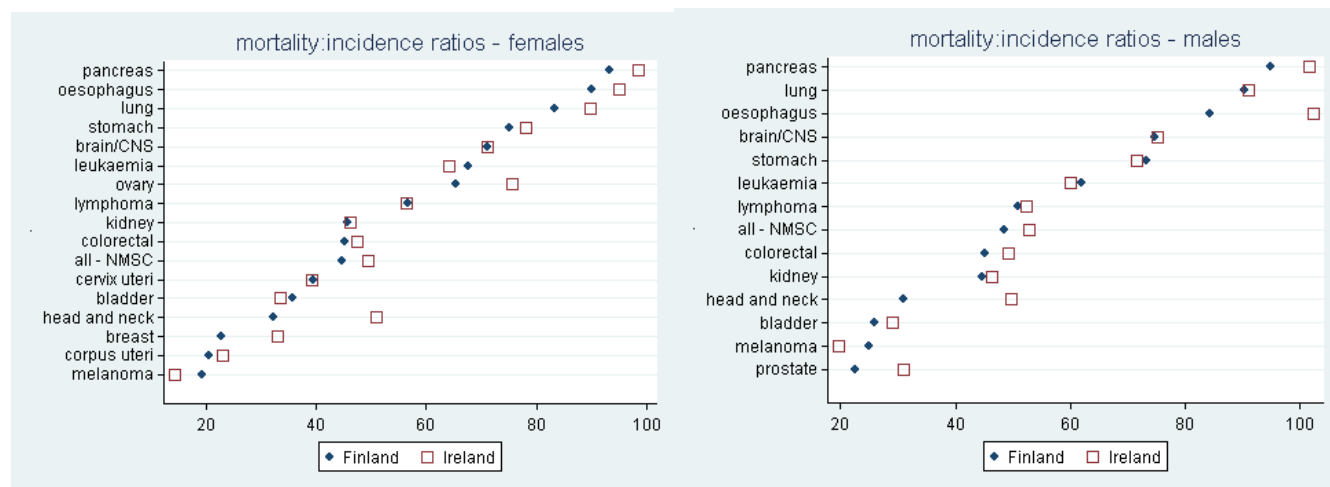
In Figure 2 three-year moving averages of age-specific incidence rates for the period 1994-2008 are shown. There are some fluctuations, with the girls in the 0-4 age-group and the boys in the 5-9 age-group showing most variation but the averages are broadly within the ranges given in Table 1.

**Figure 2: Three year moving average of incidence rates centred on middle year, childhood cancers, 1994- 2008**



Mortality/incidence ratios (M:I) for Ireland and Finland were obtained from Cancer Incidence in Five Continents, vol. IX (11) and are presented in Figure 3. In the main, Irish M:I ratios are similar to or higher than Finnish ratios. On examination of five-year survival rates (12) and (13) for both countries, we find that the differences in the M:I ratios are consistent with the differences in five-year relative survival estimates. For females, the Irish M:I ratio for head and neck is 51% and ovarian cancer is 76%; these are considerably higher than the Finnish M:I ratios of 32% and 65% respectively. For males, there is also a large difference in for head and neck with M:I ratios for Ireland and Finnish being 31% and 50% respectively. There was also a large difference in oesophagus M:I ratios for males. In both genders, Irish M:I ratios are lower for melanoma and leukaemia compared to the Finnish ratios.

Figure 3: M:I ratios for Ireland and Finland, from Cancer Incidence in Five continents, vol.IX



In Table 4 the number of sources per case is listed for all invasive cancers combined and some selected sites. While the Registry may receive many notifications regarding a particular tumour, some may be from the same source. We are interested in counting the number of distinct sources used in notifying a case. For the all non-invasive cancers (excluding NMSC), there are on average nearly three different sources received by the Registry for each case.

### Completeness : quantitative methods

#### Independent Case Ascertainment

Between 2000 and 2009, BreastCheck (6) recorded 3,926 women with a diagnosis of invasive breast cancer. Fourteen of these women had not been registered at the end of 2009 by the National Cancer Registry as having breast cancer. Eight of these 14 were ineligible for registration for a variety of reasons—non-residence at the time of diagnosis; recurrence of a previously registered cancer, or a non-registrable condition. Two of the remaining six cancers had been registered in early 2010, thus there were four cases missed by the Registry. This is a completeness of  $3,922/3,926=99.9\%$ . Completeness of breast cancer ascertainment in screening age group (50-64), using the flow method, was estimated to be 99.3%.

#### LP estimator and flow methods

The estimates for completeness using the LP estimator and the flow method are given in Table 2. The LP completeness estimate, 94.3%, for cases diagnosed in 2005 was lower than the five year completeness as estimated by the flow method, 97% (95% confidence interval, [96.2%, 97.7%]), illustrating that the LP method provides a lower bound for estimated completeness .

The flow method completeness of case ascertainment estimate increased from 87.7% after one year to 95.8% after three years. The relative increase in completeness decreases with time, and after five years was 97.0%. Estimated completeness was highest for lung cancer, and lowest for prostate cancer, a cancer with low associated mortality and low surgical intervention.



Table 2: Estimated completeness for case ascertainment, all sites combined (excluding non-melanoma skin) and four main sites, extraction date 31/12/2010

Year of diagnosis and site	ICD10 codes	flow method completeness % [95% confidence interval]			LP – lower bound for completeness %	registered while alive and died of cancer	registered while alive and did not die of cancer	not registered while alive and died of cancer	not registered and did not die of cancer (estimated)	N
		one year	three year	five year						
<b>2005 all invasive cancers excl NMSC</b>	C00-C43, C45-C96	87.7% [86.5%, 88.8%]	95.8% [95.0%, 96.6%]	97% [96.2%, 97.7%]	94.2%	37%	53%	4%	6%	15,811
<b>2005 colorectal</b>	C18-C21	91.6% [89.4%, 93.5%]	97.0% [95.8%, 98.1%]	97.4% [96.3%, 98.3%]	96.8%	41%	51%	4%	3%	1,915
<b>2005 lung</b>	C33,C34	94.6% [91.8%, 96.8%]	98.4% [96.3%, 99.7%]	98.7% [96.6%, 99.8%]	98.0%	73%	15%	10%	2%	1,590
<b>2005 breast</b>	C50	92.1% [89.1%, 94.7%]	97.1% [95.4%, 98.5%]	98.0% [96.5%, 99.1%]	95.9%	17%	78%	1%	4%	2,156
<b>2005 prostate</b>	C61	80.0% [74.7%, 83.0%]	92.8% [89.5%, 95.4%]	95.8% [93.3%, 97.8%]	90.9%	14%	76%	2%	9%	2,430

The proportion of registered cases with incomplete or unknown values for some basic variables can be a guide to the data quality. Below we show the proportion of known values for closed cases, for all cancers combined, excluding NMSC for the period 1994-2008.

Table 3: basic variable - proportion missing data

variable	Data completeness
name	100.0%
address	99.9%
gender	100.0%
date of birth	98.9%
date of death	99.8%
date of incidence	100.0%
method of diagnosis	97.3%

### 4.3 Validity

Table 4 shows the percentage of those for whom the morphology has been microscopically verified and the percentage of cases for which the only source of notification is a death certificate. Microscopic verification of pancreatic and lung cancer is low at 47.8% and 73.3% respectively, whereas Hodgkin's lymphoma and melanoma cancer have microscopic verification rates of 99.2% and 99.7% respectively. Also, the DCO rates for pancreatic and lung cancers are relatively high at 3.4% and 2.2% respectively, while Hodgkin's lymphoma has a DCO rate of 0.4% and that for melanoma is negligible.

The total percentage of cases which are listed as primary site unknown (PSU) (ICD10 code C80) in the period 2003-2007 is 2.1%. This is not homogenous across different age groups, with just 0.7% of the under 45s having a PSU listing and 1.3% and 1.9% in the 45-64 and 65-74 age groups respectively. This rises to 3.5% in the oldest age group, those aged at least 75. 91.8% of individuals with PSU died within a year of diagnosis. The DCO rate for individuals with PSU was 3%, which was high (only individuals with pancreatic cancer had a higher DCO rate at 3.4%).

The proportion of cases with missing stage information, by site is given in Table 4. The proportions vary considerably with site, from as low as 5% for breast and as high as 27% for bladder and 42% for oesophagus.

**Table 4: Number of distinct sources per case, the percentage of cases that have a specific morphology and have been microscopically (or cytology) verified and the percentage obtained from death certificates with no further information available, Ireland, 2003-2007**

site	ICD-10 code	annual average number of cases	number of sources per case	%MV*	%DCO	missing stage information
<b>all sites except NMSC</b>	C00-C96, excl C44	16146	2.8	85.4%	1.3%	-
<b>head and neck</b>	C01-C14	300	3.2	93.0%	1.1%	17%
<b>oesophagus</b>	C15	354	3.4	88.8%	1.0%	42%
<b>stomach</b>	C16	471	3.0	90.0%	1.8%	25%
<b>colorectal</b>	C18-C21	2152	2.8	91.5%	0.9%	10%
<b>pancreas</b>	C25	415	3.0	47.8%	3.4%	22%
<b>lung</b>	C34-C34	1866	3.3	73.7%	2.2%	14%
<b>melanoma</b>	C43	615	2.2	99.7%	0.0%	10%
<b>breast</b>	C50	2289	2.9	95.0%	0.6%	5%
<b>cervix uteri</b>	C53	239	3.1	95.8%	0.4%	9%
<b>corpus uteri</b>	C54	293	2.8	95.5%	0.7%	9%
<b>ovary</b>	C56	342	2.8	81.8%	2.0%	9%
<b>prostate</b>	C61	2498	2.6	90.9%	1.0%	20%
<b>kidney</b>	C64-C66, C68	420	2.7	74.0%	1.4%	7%
<b>bladder</b>	C67	472	2.6	90.8%	0.7%	27%
<b>brain/CNS</b>	C70-C72	313	3.4	68.3%	2.0%	-
<b>Hodgkin's lymphoma</b>	C81	103	2.6	99.2%	0.4%	5%
<b>non-Hodgkin's lymphoma</b>	C82-C85, C96	824	2.8	83.6%	0.5%	13%
<b>myeloma</b>	C90	210	3.0	89.8%	1.6%	-
<b>leukaemia</b>	C91-C95	469	2.6	94.2%	1.7%	-

\*Microscopically verified with specific morphologies as a percentage of all invasive tumours for that site, where non-specific morphologies are M-8000, M-8001, M-8010, for haematology non-specific morphologies are M-9800, M-9590.

In Figure 4, we show bar charts comparing the percentage of cases which are microscopically verified in Ireland with other European countries (11).

Figure 4: Percentage of cases morphologically verified, all cancer sites combined, excluding NMSC



## 4 Discussion

The historical time trends are consistent with steady case reporting over time, while the childhood cancer rates for the total period (1994-2008) are within the limits of the reference interval. Childhood cancer rates show the stability in incidence rates which would be expected if registration practices and completeness were consistent over the period of operation of the registry.

We compared the Irish M:I ratios for various cancer sites with Finnish M:I ratios as Finland is known to have high quality data with compulsory reporting of cancer since 1961. The mortality incidence ratios for Finland tend to be lower than for Ireland. Some cancer sites are showing large differences in M:I ratios.. In particular, the M:I ratios for Ireland for cancer of the head and neck is higher for both males and females than the Finnish ratios. A higher M:I ratio may be due to under-registration of cases, over-registration of deaths (due to misclassification of the cause of death) or higher case fatality. These cannot usually

be distinguished; however survival for most cancers is higher in Finland, which may explain much of the difference in M:I ratios. In the case of head and neck cancer, death certificates may be inaccurate in the anatomical localisation of the cancer, leading to misclassification.

Registration of cases from more than one independent source makes the possibility of missing a case much less. An average of 2.8 sources per case in Ireland gives some reassurance that few cases are likely to be missed.

The quantitative methods indicate that case ascertainment in the Irish National Cancer Registry is at approximately 97% at five years from the year of diagnosis. Overall completeness of registration (from the flow method) was 95.8% at three years from diagnosis and 97.0% at five years, and was above 97% for three of the four commonest cancer sites. There was close agreement between the two methods of estimation for lung cancer, since its high mortality rate enables the LP bound to be tighter for this site.

We consider the estimated completeness levels of Norway, which has a comparable population size to the Republic of Ireland. Norway has a longer history of cancer registration, over 50 years and also has mandatory reporting of cancer. The 5-year completeness of 97.0% of the Irish National Cancer Registry was close to that reported by the Norwegian Cancer Registry for cases diagnosed in 1999 of 97.8% (14).

Of the four cancer sites considered, completeness estimates were lowest for prostate cancer. The lower bound provided by the LP estimator of 91% was much lower than that estimated by the flow method with a five year estimate of 96%. This indicates that while prostate cases are eventually registered, registration is slow. This may be due to the fact that, compared to most cancers, rates of surgery are low, and a significant proportion of cases are treated initially by androgen deprivation therapy (ADT) or active surveillance. Typically men on ADT are issued an initial prescription by a urologist and then are re-issued prescriptions in primary care, making them more difficult cases to identify. Men on active surveillance have prostate specific antigen tests regularly in primary care, and may have prostatic biopsies intermittently. This means that the patients leave little footprint in hospital systems and records in the early phases of treatment and are hence difficult for a registry to ascertain.

The flow method and the LP estimator rely on various assumptions in order to estimate the completeness of cancer registration. Both methods rely on key parameters such as incidence and mortality rates being in a steady state and accurate specification of cause of death on death certificates. The LP estimator relies on the assumption that the case fatality in those registered while alive is the same as that in those registered through death certificate notification. The lower bounds for completeness using the LP estimator are particularly low for breast and prostate cancer and high for lung cancer. This may be due to the lower mortality rates for these breast and prostate cancer and higher mortality for lung cancer. In addition, the estimated completeness is based on the extracted date of 31/12/2010 so that cases have on average 5.5 years follow-up. For these cancers, one can still expect that further cancer deaths will occur and thus we would expect that the M:I ratio to increase further, before reaching an equilibrium level.

Fundamentally, both methods simply provide an *estimate* of levels of missing data; the true level of completeness cannot, by definition, be computed. Two studies (15; 16) have compared the LP estimator (also known as the DCN/Ajiki method (17; 18)) and flow methods to “true” completeness using simulated data. Silcocks et al. (15) showed that a naive LP estimator grossly underestimated completeness, while Schmidtmann (16) concluded that a version of the LP estimator provides the least biased estimator in most situations. A better estimate of the mortality to incidence ratio for registered patients would improve the LP estimator and Schmidtmann (16) has suggested that taking the number of deaths occurring in the year in question, rather than the number of deaths due to those diagnosed in the year in question would give an improved result.

Independent case ascertainment performed for screen-detected breast cancer gave a higher estimate of completeness than the overall LP or flow methods, but was reasonably close to the estimate for the specific age-group and cancer site studied. Women with screen-detected cancers are aged between 50 and 65 and almost all have surgical treatment, so these cancers will leave a significant footprint in hospital systems and records and are therefore very likely to have been registered. However, completeness in this case was so high that, although the correspondence between the statistical estimates and independent verification is reassuring, this should not be taken as an overall verification of the estimation approach.

The indicators used to assess validity were the percentage of cases which were morphologically verified or were listed as primary site unknown. We also examined the percentage of cases with stage unknown. Although the Registry TROs always attempt to assign a TNM category from the clinical information, this is not always possible. Information may be incomplete because staging investigations were not carried out, or because the information, although known to the clinician, was not recorded in the medical notes. Stage information was missing in a significant number of cases, particularly for cancers where surgery was less likely to be carried out. Morphologically verified cases reached 85.5% for Irish cases in the period 2003-2007. This figure seems reasonable, especially as a very high number here may signal an over-reliance on pathology reports for case ascertainment. The overall DCO rate is 1.3% which is fairly low. This is at the median level for 27 European registries (data not shown).

The registry complies with international standards and criteria in recording of all major data items, and has been consistent in its application of these since the start of registration in 1994. Overall, the results suggest that the completeness, validity and timeliness of the data at the NCR are of a good standard, though there is some heterogeneity between cancer sites.

## Bibliography

1. *Evaluation of data quality in the cancer registry: Principles and methods. Part1. Comparability, Validity and Timeliness.* **Bray F, Parkin DM.** 2009, Eur J Cancer, Vol. 45, pp. 747-755.
2. *Evaluation of the data quality in the cancer registry: Principles and methods. Part II. Completeness.* **Parkin DM, Bray F.** 2009, Eur J Cancer, Vol. 45, pp. 756-764.
3. *Comparability and Quality Control in Cancer Registration (IARC Technical Report No. 19).* **Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL.** Lyon : IARC, 1994.

4. "Calculating Waterfowl Abundance on the Basis of Banding Returns". **Lincoln, FC.** 1930, United States Department of Agriculture Circular, Vols. 1–4, p. 118.
5. *Completeness of cancer registration : a new method for routine use.* **Bullard J, Coleman MP, Robinson D, Lutz J-M, Bell J, Peto J.** 2000, British Journal of Cancer, Vol. 82, pp. 1111-1116.
6. Breastcheck in Ireland. [Online] [Cited: 19 April 2011.] <http://www.breastcheck.ie/>.
7. *International Rules for multiple primary cancers.* **International Association of Cancer Registries.** 2005, Asian Pac J Cancer Prev, Vol. 6, pp. 104-6.
8. *International rules for multiple primary cancers (ICD-0 third edition).* **Working Group Report.** 2005, Eur J Cancer Prev, Vol. 14, pp. 307-8.
9. **Fleming ID, Cooper JS, Murphy GP, Sullivan BO, Sobin LH, Yarbrow JW, et al.** *American Joint Committee on Cancer: AJCC Cancer Staging Manual.* 5th edition. Philadelphia, PA, USA : Lippincott-Raven, 1997.
10. **Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB, editors.** *Cancer incidence in five continents, vol VIII (IARC Scientific Publications No. 155).* Lyon : IARC, 2002. pp. 57-73.
11. **Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M and Boyle P.** *Cancer incidence in five continents, vol IX (IARC Scientific Publications No. 160).* Lyon : IARC, 2007.
12. Newest survival ratios - Cancer Society of Finland. [Online] [Cited: 15 Decemeber 2011.] <http://www.cancer.fi/syoparekisteri/en/statistics/newest-survival-ratios/> .
13. The National Cancer Registry, Ireland. [Online] [Cited: 14 December 2011.] [http://www.ncri.ie/survival/RELATIVESURVIVAL\\_WEBPAGE.shtml](http://www.ncri.ie/survival/RELATIVESURVIVAL_WEBPAGE.shtml) .
14. *Data quality at the Cancer Reigstry of Norway: An overview of comparability, completeness, validity and timeliness.* **Larsen IK, Småstuen M, Johannesen T, Langmark F, Parkin DM, Bray F, Møller B.** 7, 2009, European Journal of Cancer, Vol. 45, pp. 1218-31.
15. *Simulation modelling to validate the flow method for estimating completeness of case ascertainment by cancer registries.* **Silcocks PB, Robinson D.** 2007, J Public Health, Vol. 29, pp. 455-62.
16. *Estimating Completeness in Cancer Registries - Comparing Capture-Recapture Methods in a Simulation Study.* **Schmidtman I.** 6, 2008, Biometrical Journal, Vol. 50, pp. 1077-92.
17. *Index for evaluating completeness of registration in population-based cancer registries and estimation of registration rate at the Osaka Cancer Registry between 1966 and 1992 using this index.* **Ajiki W, Tsukuma H, Oshima A.** 1998, Nippon Koshu Eisei Zasshi, Vol. 45, pp. 1011-7. (in Japanese, English abstract).
18. *A mathematical estimation of true cancer incidence using data from population-based cancer registries.* **Kamo K, Kaneko S, Satoh K, Yanagihara H, Mizuno S and Sobue T.** 2007, Jpn J Clin Oncol, Vol. 37 (2), pp. 150-155.