1. INTRODUCTION

The cancer registries in Northern Ireland and the Republic of Ireland became operational almost simultaneously in 1993-1994, and quickly built up a strong collaborative relationship. The collaboration has been manifest in three all-Ireland cancer incidence reports, numerous joint peer-reviewed publications, several joint research projects and extensive informal exchanges of information, expertise and support. This relationship was recognised at the establishment of the NCI-Ireland-Northern Ireland Cancer Consortium by the creation of a Cancer Registries Group, later to become the Cancer Registries and Epidemiology Group. Both registries realise that, on a small island, there is much potential in working together and in sharing our links with the UK and the rest of Europe.

This atlas is the latest fruit of our collaboration. Building on earlier work which produced an atlas of cancer in the Republic of Ireland, this document has been produced by a team of statisticians, epidemiologists and GIS specialists on both sides of the border. Carrying out the analysis has posed some challenges because of the different nature of the small geographical areas available for analysis and the incompatibility of almost all census measures of socio-economic status between the two countries. However, despite these limitations, this atlas gives, for the first time, an overview of the distribution of cancer risk across the island of Ireland.

AIMS

The aims of this atlas were:

1. to describe, through mapping and statistical analysis, geographical variation in cancer risk at small area level (electoral division and ward) across the island of Ireland;
2. to describe socio-economic and demographic effects on cancer risk;
3. to attempt to relate the variations found to the distribution of known cancer risk factors and other determinants of health;
4. to make recommendations based on our findings.

BACKGROUND

Geographical variation in cancer risk has been a subject of fruitful research since cancer registration began in the 1940s. Although often posing questions rather than providing answers, inspection of geographical patterns and their variation between cancer types and over time can give valuable indications of the likely cancer risks and their distribution in our populations. Sometimes, too, examination of this variation can lead to the discovery of new aetiological agents. Within a relatively homogeneous population such as Ireland's, genetic variation is an unlikely reason for geographical differences in cancer risk, and most must be attributed to modifiable factors of some sort. The majority of variation in modifiable cancer burden is known to be due to four "lifestyle" factors—tobacco, diet, alcohol and sexual/reproductive life (Doll and Peto, 1981). Much less of the risk appears to be attributable to what are commonly called "environmental" factors—radiation exposure, carcinogens in water, air and food, and other external causes. In explaining the variations seen, we therefore need to look closely at people as well as places, and in this atlas we have provided analyses of the personal characteristics, insofar as we could measure them, of people living in low- and high-risk areas.

WHAT THIS ATLAS CONTAINS

This atlas describes the geographical distribution of the 18 commonest cancers in two ways—through statistical analysis of the variation of cancer risk by characteristics of small areas of residence (electoral division in the Republic of Ireland and ward in Northern Ireland) and in smoothed maps of relative risk for the

small geographical units. Each cancer has been assigned a separate chapter, which gives a general summary of the incidence data for the cancer, international comparative incidence rate, risk factors, analysis by small area characteristics and smoothed maps of relative risk. The chapters are ordered by the frequency of occurrence of the cancers—most frequent first.

The statistical analyses describe models of cancer risk at small area level based on the socio-demographic characteristics of the areas—unemployment, education, rurality etc. Implicit in the models is the hypothesis that the characteristics of the small areas are related in some way to those of the individuals living in them. We have explored the strengths and limitations of this approach (Cook et al., 2000) in the previous atlas of the Republic of Ireland (Carsin et al, 2009).

Mapping of cancer risk at small area level, particularly when smoothing techniques are used, as they are in this atlas, is based on the assumption that populations living in adjacent areas are likely to share the same risk factors, and therefore have similar underlying risks of cancer. This hypothesis suggests that most of the variation in cancer incidence between small areas is due to random variation and that by smoothing this variation over larger areas we can arrive at a better estimate of true cancer risk. However, while maps may give a valuable overview of cancer distribution, in one respect they may be deceptive. The majority of the Irish population, north and south, is urban, and living in a few cities of relatively small area, which are inconspicuous on the maps. On the other hand, large areas of the country, particularly in the west and north-west, are sparsely populated but very visible on the map. As a result, the appearance of the maps tends to be dominated by risk in the rural population, while the analyses by socio-economic and demographic area characteristics mainly represent the urban population. The two approaches are complementary.

We hope that this atlas is a beginning, rather than just an end in itself, and that the many questions raised by our analyses will stimulate studies which can make a real impact on understanding, and ultimately reducing, cancer risk in Ireland.

2. METHODS

2.1 GEOGRAPHICAL TERMS

From the Republic of Ireland perspective, the island of Ireland, and the country comprising most of its area, are both officially known as “Ireland”, while in Northern Ireland there is no universally agreed terminology for the different geographic bodies in the island of Ireland. As this is certain to cause confusion in an atlas of this kind, we have elected to use the expressions “Republic of Ireland (RoI)” and “Northern Ireland (NI)” for the two jurisdictions on the island and to refer to both of these areas as “countries”. The combined area of the whole island (and the offshore islands) is referred to in the text simply as “Ireland”. None of this implies anything concerning the official status of these names.

For administrative purposes, RoI is divided into 27 counties (including Tipperary North and South) (Map 2.1). Counties with large urban areas are further divided into “city” and “county” areas (three of the latter in the case of Dublin), giving a total of 34 large administrative areas. Small area population statistics are available at the level of electoral division (ED), of which there are approximately 3,500.

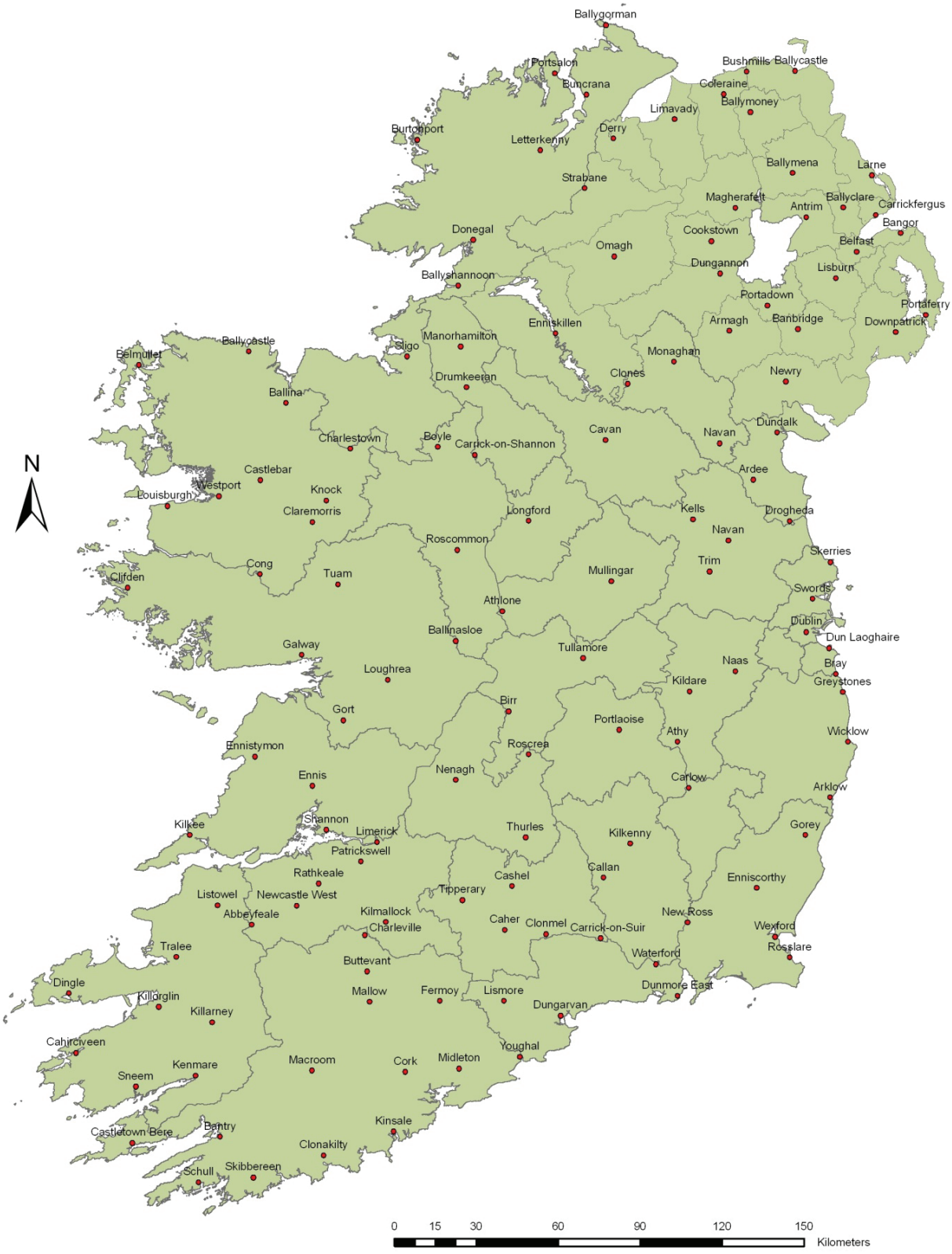
NI has six counties and is also divided into 26 district councils—four city councils (Armagh, Belfast, Derry and Lisburn), 13 Borough Councils and 9 District Councils (Map 2.1). Small area population statistics are available at the level of ward, of which there are approximately 580. District councils in NI therefore have smaller populations on average than counties in RoI, while NI wards have an average population size greater than RoI electoral divisions (see section 2.2.4). Mapping and statistical analysis in this atlas is based almost exclusively on data at the ward/ED level.

A number of other geographical entities are also referred to in the atlas—these include health board areas, health service regions, provinces and planning regions. These are described in Appendix table A4.1.

Map 2.2 shows, for reference, the outlines of counties in RoI and District Councils in NI with some towns and cities.

Map 2.1 Counties and district councils

Map 2.2 Locations



2.2 DATA INCLUDED IN THE ATLAS

2.2.1 CANCER REGISTRATIONS

NORTHERN IRELAND

The Northern Ireland Cancer Registry (NICR) was established in 1994 and uses an automated computer system with multiple information sources to collate information on new diagnoses of cancer. Information has been collected for incidence years from 1993 onwards. The three main sources for registration are the Patient Administration System (PAS) used by all the Hospital Trusts, histopathology reports, and death notifications supplied by the General Register Office (GRO). From PAS the Registry obtains demographic information on individual patients along with basic site and behaviour information for each tumour. This information is supplemented by electronic downloads from histopathology and cytopathology laboratories. A major focus of the Registry's operation is on the verification of the information from a single hospital admission, a single histopathology report or a single death certificate (death-certificate initiated cases). Trained Tumour Verification Officers (TVOs) examine general practitioners' (GPs) notes for patients who have died from cancer, hospital records for cases identified without histopathology or cytology confirmation and histopathology reports where there is conflicting information or other possible errors. In the event that no further information on death-certificate initiated cases is obtainable the record is included in the Registry but flagged as a death certificate only (DCO) case. These comprised less than 2% of cases in 1995-2007. Follow up of patients is conducted passively by linking cancer incidence data to death certificate information.

In the NICR the information on cancer site received by the Registry has been coded using the Systematized Nomenclature of Medicine (SNOMED) which is used in the UK National Health Service (NHS) and information on cancer morphology has been coded to the second revision of the International Classification of Diseases for Oncology (ICD-O-2) (World Health Organisation, 1990). Cancer site is recoded at the Registry to the tenth revision of the International Classification of Diseases (ICD-10) (World Health Organisation, 1997).

REPUBLIC OF IRELAND

The National Cancer Registry Ireland (NCRI) was established in 1991 and has produced national figures on cancer incidence since 1994. Most registrations are based on "active" data collection, whereby trained Tumour Registration Officers (TROs), based in hospitals around the country, access a range of data sources to identify all new cancer cases and register all relevant patient, tumour and treatment details. Hospital pathology reports, provided to the Registry shortly after diagnosis, comprise the bulk of information, providing data on approximately 85% of all new cases. Most pathology reports are registered manually by the TROs but about 10% of pathology reports are now provided in electronic format. Information on non-microscopically diagnosed cases is registered mainly from other hospital sources, principally the Hospital Inpatient Enquiry system (HIPE) as well as records from radiology and oncology departments, medical charts, etc. Most cases (≥95%) are registered in this way. The main non-hospital source of case information is death certificate data. The Registry is provided with all death certificates by the Central Statistics Office (CSO). All cases initially notified by death certificate are followed up with the hospital of death or the certifying doctor and most cases are subsequently found in other data sources. Only a small percentage of cases (<3%) remain classified as notified by death certificate only (DCO). As in the NICR, follow up of patients is passive, where cancer cases are linked to death certificate information provided regularly from the CSO.

Although case data from pathology reports is registered almost immediately after diagnosis, data from other sources can take longer to obtain. Together with essential case checking and data quality assurance, the Registry normally produces definitive statistics for case data a minimum of 18 months to 2 years following the end of year of diagnosis. Currently the completeness of cancer registration for all invasive cancers diagnosed to end 2007 is estimated to be over 96%.

METHODS

Incident cases are coded according to the third edition of the International Classification of Diseases for Oncology (ICD-O-3) (World Health Organisation, 2000).

2.2.2 ATLAS DATA

Before analysis, data on cases diagnosed 1995-2007 were recoded to the equivalent ICD-10 classification in both registries. Cases were extracted from both registry datasets, based on ICD-10 codes, and data from Northern Ireland (NICR) and Republic of Ireland (NCRI) were then amalgamated into a single dataset. Table 2.1 lists the 18 cancer sites analysed, with the relevant ICD-10 codes and total number of cancers in RoI and NI. Data from both registries have previously been used in three joint all-Ireland incidence reports—further information on data comparability and quality is provided in Donnelly et al., 2009. Multiple primary cancers were excluded in the calculation of incidence figures, based upon the rules published by the International Agency for Research on Cancer (IARC) (Ferlay et al., 2005). Death certificate only registrations were also excluded. For all cancers, maps are based on the incidence period 1995-2007. For breast and prostate cancers additional analysis was carried out, and maps produced, for the periods 1995-2001 and 2002-2007 separately because of the introduction of mammographic screening in RoI in 2000, and striking temporal trends in both countries in both prostate cancer incidence and prostate-specific antigen testing (Carsin et al., 2010).

Table 2.1 Incident cancers diagnosed 1995-2007 included in this report: Republic of Ireland (RoI) and Northern Ireland (NI)

cancer site	ICD-10 codes	Ireland		RoI		NI	
		females	males	females	males	females	males
non-melanoma skin cancer	C44	49,102	55,826	34,655	40,034	14,447	15,792
breast	C50	38,545	*254	25,876	*181	12,669	*73
colorectal	C18-C21	16,992	21,205	11,041	14,485	5,951	6,720
lung	C34	13,005	20,831	8,437	13,672	4,568	7,159
prostate	C61	—	33,144	—	24,704	—	8,440
non-Hodgkin's lymphoma	C82-C85	4,605	5,094	2,917	3,441	1,688	1,653
stomach	C16	3,616	5,748	2,353	3,818	1,263	1,930
melanoma of the skin	C43	5,467	3,702	3,871	2,611	1,596	1,091
bladder	C67	2,513	6,226	1,730	4,309	783	1,917
head and neck	C01-C14,C30-C32	2,211	5,692	1,371	3,820	840	1,872
leukaemia	C91-C95	3,164	4,527	2,235	3,331	929	1,196
pancreas	C25	3,533	3,502	2,499	2,454	1,034	1,048
kidney	C64-C65	2,450	4,032	1,603	2,785	847	1,247
oesophagus	C15	2,370	3,907	1,592	2,627	778	1,280
ovary	C56	6,222	—	4,149	—	2,073	—
brain and other central nervous system	C70-C72	2,266	3,041	1,630	2,161	636	880
corpus uteri	C54	5,237	—	3,355	—	1,882	—
cervix uteri	C53	3,758	—	2,665	—	1,093	—

** Since breast cancer in males is rare, most of the analyses in chapter 4 are limited to breast cancer in women.*

2.2.3 GEOCODING OF CANCER CASES TO ELECTORAL WARDS (NI) AND ELECTORAL DISTRICTS (ROI)

Apart from its obvious use in mapping, geocoding of cancer cases allows linkage to area-based data, such as population density and measures of socio-economic status (e.g. percentage unemployed). This is described in more detail in sections 2.2.4.2 and 2.2.4.3. This type of information is not, in general, accessible at the level of the individual cancer case in Ireland, and has to be inferred from area-based measures.

NORTHERN IRELAND

NICR routinely collects address information for registered cancers, allowing small geographic areas to be assigned to individual cancer registrations. This is accomplished through an electronic process which uses the postcode that accompanies the majority of NI addresses along with a postcode-to-electoral ward lookup file

known as the Central Postcode Directory (CPD). This is maintained by the Northern Ireland Statistics and Research Agency (NISRA) and updated annually (Northern Ireland Statistics and Research Agency, 2010a). Addresses with an unknown, incomplete or invalid postcode cannot be assigned an electoral ward.

REPUBLIC OF IRELAND

The National Cancer Registry attempts to code all addresses of cancer cases to the level of electoral division (ED), the smallest area for which census data can be obtained. Unlike NI, addresses in RoI do not have postcodes. Each address is therefore assigned to an ED by means of matching Registry address information to other data sources. This process of geocoding is carried out for the most part by matching Registry patient address data to the GeoDirectory database which provides a list of official postal addresses and location details for every residential and commercial property in the country (www.geodirectory.ie). Data matching is carried out using software developed by the Registry, with some manual coding for any remaining unmatched records. Additional resources used include address tables from census surveys, supplied by the Central Statistics Office (CSO), as well as manually locating addresses on maps provided by Ordnance Survey Ireland (OSI). Using a combination of these resources, almost all complete patient addresses can be assigned to a particular ED.

For some cases it is impossible to assign an address confidently to a single ED, usually because the address is incomplete or ambiguous (4.4% of all RoI cancers included in this report; Table 2.2). The number of EDs to which the address could potentially belong is usually small (2 or 3) and, for analysis, these cases were assigned at random to one of the possible EDs, with the possibility of assignment weighted by the population of the alternative EDs.

NI AND ROI RECORDS WITH UNKNOWN ED OR WARD

At the end of the geocoding process, a number of registrations in both RoI and NI remained which could not be assigned to any ED/ward (3.3% of all cancers included in this report; 3.6% in RoI and 2.7% in NI) (Table 2.2). For these registrations, a fraction of the cases of each cancer type was allocated in proportion to each ED (RoI) or ward (NI) weighted by population. In NI almost all of these cases were non-melanoma skin cancers, but in RoI all cancer sites had a similar percentage of cases with unknown ED.

Table 2.2 Number and percentage of cases not assigned to an ED or ward, and of cases assigned to multiple EDs

cancer site	cases assigned to more than one ED (RoI only)			cancer cases not assigned to an ED or ward					
				Ireland		RoI		NI	
	number	% RoI cases	% all cases	number	%	number	%	number	%
non-melanoma skin cancer	3443	4.6%	3.3%	4896	4.7%	2803	3.8%	2093	6.9%
breast	1045	4.0%	2.7%	948	2.4%	839	3.2%	109	0.9%
colorectal	1136	4.5%	3.0%	985	2.6%	885	3.5%	100	0.8%
lung	783	3.5%	2.3%	808	2.4%	749	3.4%	59	0.5%
prostate	1282	5.2%	3.9%	1128	3.4%	983	4.0%	145	1.7%
non-Hodgkin's lymphoma	270	4.2%	2.8%	305	3.1%	254	4.0%	51	1.5%
stomach	256	4.1%	2.7%	270	2.9%	250	4.1%	20	0.6%
melanoma of the skin	280	4.3%	3.1%	365	4.0%	281	4.3%	84	3.1%
bladder	248	4.1%	2.8%	222	2.5%	208	3.4%	14	0.5%
head and neck	199	3.8%	2.5%	190	2.4%	168	3.2%	22	0.8%
leukaemia	258	4.6%	3.4%	247	3.2%	230	4.1%	17	0.8%
pancreas	243	4.9%	3.5%	173	2.5%	164	3.3%	9	0.4%
kidney	205	4.7%	3.2%	129	2.0%	120	2.7%	9	0.4%
oesophagus	203	4.8%	3.2%	169	2.7%	151	3.6%	18	0.9%
ovary	183	4.4%	2.9%	165	2.7%	154	3.7%	11	0.5%
brain and other central nervous system	168	4.4%	3.2%	155	2.9%	150	4.0%	5	0.3%
corpus uteri	129	3.8%	2.5%	133	2.5%	113	3.4%	20	1.1%
cervix uteri	75	2.8%	2.0%	90	2.4%	81	3.0%	9	0.8%
all cancers in this report	10406	4.4%	3.0%	11378	3.3%	8583	3.6%	2795	2.7%

2.2.4 CHARACTERISTICS OF EDS AND WARDS: POPULATION AND SOCIO-ECONOMIC VARIABLES

2.2.4.1 POPULATION

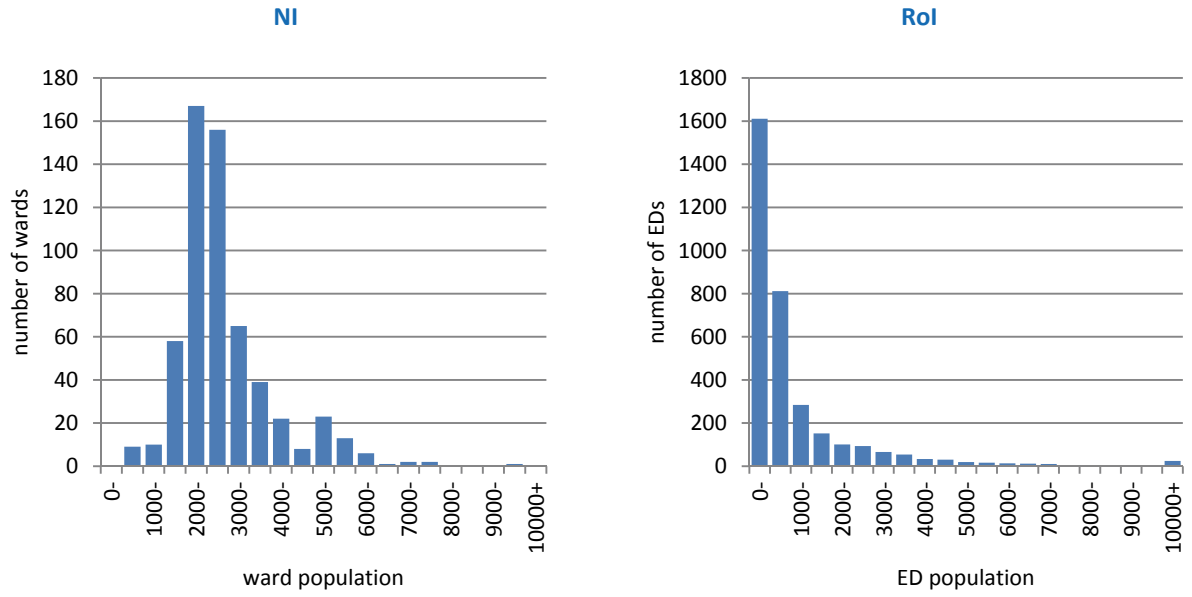
NORTHERN IRELAND

A census of population was carried out in NI in 2001, the only census between 1995 and 2007. This census provided population data, broken down by sex and age, for 582 wards in 26 district councils. Population estimates for each year were available by sex and age at district council level. Annual estimates for the wards were derived from these total annual estimates, using the 2001 census as the basis for the splits by ward. The estimates for each year were then averaged to give an estimated average population by ward for the 1995-2007 period. Over this period the wards had an average population of 2,913, ranging from 784 (Bushmills, Moyle) to 9,654 (Botanic, Belfast) (Table 2.3, Figure 2.1).

Table 2.3 Population distribution of NI wards and RoI EDs

	number of areas	mean population	standard error of mean	standard deviation	minimum population	25th percentile	median	75th percentile	maximum population
NI wards	582	2913	48	1147	784	2219	2618	3238	9654
RoI EDs	3355	1161	34	1956	62	309	525	1146	33983

Figure 2.1 Population distribution of NI wards and RoI EDs



REPUBLIC OF IRELAND

Three censuses were carried out in RoI during the period of this report, in 1996, 2002 and 2006. These censuses provided population data, broken down by sex and age, for 3,422 EDs in 1996 and 2002, and for 3,409 EDs in 2006. Population data were derived from the census small area population statistics (SAPS) files for 1996, 2002 and 2006. Official CSO estimates of the total population split by sex and age (but not by ED) were available for each year from 1995 to 2007. Annual estimates for the EDs were derived from the appropriate census and the CSO total annual estimates–the 1996 census results were used as the basis for the ED populations for 1995, a linear interpolation of the 1996 and 2002 census counts was used for 1997-2001, a linear interpolation of the 2002 and 2006 census counts was used for 2003-2005, and the 2006 census results

were used for 2007 estimates. The estimates for each year were then averaged to give an estimated average population by ED for the period 1995-2007.

The average ED population over the period was 1,161; ranging from 62 (Mountstuart, Co. Waterford) to 33,983 (Dundalk Urban, Co. Louth) (Table 2.3, Figure 2.1). Dundalk Urban District comprised a number of EDs in 2006 which were merged for the purposes of this atlas (see below); this merged area was the largest single population unit treated as an ED in RoI. The population of the largest single ED (Blanchardstown-Blakestown) was 23,179.

At each census, the population of a number of EDs was so low that the CSO considered these EDs "confidential", published only total population figures for them, and amalgamated them with one or more neighbouring EDs for the purpose of reporting age-specific population numbers. EDs were considered confidential by the CSO if they included either 15 households or less, or 50 persons or less. There were 12 such confidential EDs in 1996, 19 in 2002 and 32 in 2006. Three of the 2006 confidential EDs had been merged with different EDs in 2002 and so, to create an estimated population for each ED for 1995-2007, any EDs that had been merged during any of these censuses were combined. These are shown in Appendix table A2.1.

The definition of a small number of EDs, and therefore the associated SAPS data, changed between the 1996 and 2002 censuses. These changes consisted of splitting or amalgamation of areas, rather than any movement of boundaries. EDs which had changed in this way were combined for analysis, and the available age and sex distribution similarly combined (Appendix table A2.2). In addition, between 1996 and 2006 there was considerable population growth in a number of towns, many of which consisted of a single ED (urban part), with a surrounding ED (rural part). As the population of these towns increased, they expanded into the rural area, but the ED boundaries remained unchanged. Because of the uncertainty of geocoding of new buildings in these towns, the urban and rural EDs were combined for analysis (Appendix table A2.3). Finally, for the towns of Drogheda, Dundalk and Wexford, population splits were not available for all EDs for all censuses, and the affected EDs were also merged for analysis (Appendix table A2.4). The population of the largest merged ED (Dundalk Urban) was 33,983. This combining of areas gave a final total of 3,355 EDs.

2.2.4.2 POPULATION DENSITY

As the formal definition of “urban” areas in Ireland (RoI and NI) does not include many areas at the periphery of towns and cities, urban and rural populations were distinguished by population density (Table 2.4), based on the estimated average number of inhabitants in 1995-2007. Three categories were created for analysis, with the cut-off points (<1 person/hectare, 1-15 persons/hectare, >15 persons/hectare) chosen to give an approximately equal population in each group.

Table 2.4 Distribution of cancer cases and estimated average population in 1995-2007, and number of EDs and wards, by population density tertiles

population density	no. of cancer cases*	estimated average population	% of total population	number of EDs and wards
<1 person/ha	121,810	2,004,451	36%	2,892
1-15 persons/ha	90,597	1,644,792	29%	403
>15 persons/ha	129,380	1,940,844	35%	642
Total	341,787	5,590,087		3,937

** All cancers included in this report.*

2.2.4.3 SOCIO-ECONOMIC INDICATORS

A range of area-based socio-economic measures is available from the population censuses in NI and RoI. However, the majority of these, particularly those relating to occupation and social class, use different definitions in NI and RoI, and are not directly comparable. Three measures were identified as having a degree of compatibility and have been used for analysis in this report:

1. Unemployment–the proportion of the economically active population aged 16-74 who were unemployed (based upon the definition of unemployment[1] from the International Labour Office (ILO))
2. Educational attainment–the proportion of people aged 16-74 who had a university degree. Academic qualifications which were equivalent to a university degree were included; however, professional qualifications were not, as this information was not available in NI.
3. Elderly living alone—the proportion of people aged 75 and older who lived alone.

These socio-economic measures had to be changed from those in the RoI cancer atlas (Carsin et al., 2009), as the necessary information was not available from the 2001 NI census.

Wards and EDs were ranked according to increasing levels of each of these three variables and were divided into population quintiles, (i.e. each quintile contained as close to 20% of the population as possible). The 20% of the population resident in areas with the lowest percentage of, for instance, unemployment, was assigned to quintile 1 while the 20% resident in areas with the highest percentage was assigned to quintile 5. All measures were based upon data for men and women combined from the censuses of 2001 in NI and 2002 in RoI.

VARIATIONS BY COUNTRY

Overall, 40% of the NI 16-74 year old population was economically inactive compared to 34% in RoI. Of the economically active population 7% in NI were unemployed in the 2001 census compared to 8% in the 2002 RoI census (Northern Ireland Statistics and Research Agency, 2003; Central Statistics Office, 2003). While, overall, 20% of the population of the island was resident in each unemployment quintile, 30% of the NI population lived in the areas of highest unemployment, compared to 16% of the RoI population (Table 2.5).

Among 16-74 year olds in RoI, 87% did not have a university degree (or academic equivalent) compared to 84% in NI. 25% of the RoI population lived in the areas with the lowest level of tertiary-level education in Ireland, compared to 10% of the NI population.

41% of the NI population aged 75 years and over lived alone, compared to 31% in RoI. 39% of the NI population lived in areas with the highest level of elderly living alone, compared to 12% of the RoI population.

[1] *An unemployed person is a person who is not in employment, is available to start work in the next 2 weeks, and has either looked for work in the last 4 weeks or is waiting to start a new job.*

Table 2.5 Population and number of areas (wards and EDs) included in each area-based socio-economic category

		RoI			NI			Ireland		
	Quintile Range	Number of areas	Population*	% of total population	Number of areas	Population*	% of total population	Number of areas	Population*	% of total population
Unemployment; % of economically active persons, aged 16-74, who were unemployed										
Least unemployed (Q1)	0.0% - 3.5%	967	754,815	19%	114	358,345	21%	1,081	1,113,160	20%
Quintile 2	3.6% - 4.7%	695	894,496	23%	78	225,691	13%	773	1,120,187	20%
Quintile 3	4.8% - 6.2%	672	831,031	21%	104	283,577	17%	776	1,114,608	20%
Quintile 4	6.3% - 8.6%	593	803,616	21%	121	320,098	19%	714	1,123,713	20%
Most unemployed (Q5)	8.7% - 47.3%	428	610,592	16%	165	507,827	30%	593	1,118,419	20%
Total		**3,355**	**3,894,549**		**582**	**1,695,538**		**3,937**	**5,590,087**	
Education; % of persons aged 16-74 without a university degree (or academic equivalent)										
Least with no degree (Q1)	44.5% - 81.0%	239	691,541	18%	123	425,072	25%	362	1,116,612	20%
Quintile 2	81.1% - 86.2%	319	682,943	18%	146	434,777	26%	465	1,117,720	20%
Quintile 3	86.3% - 89.5%	557	714,096	18%	149	402,893	24%	706	1,116,989	20%
Quintile 4	89.6% - 92.7%	978	847,567	22%	110	271,190	16%	1,088	1,118,757	20%
Most with no degree (Q5)	92.8% - 100.0%	1,262	958,403	25%	54	161,606	10%	1,316	1,120,009	20%
Total		**3,355**	**3,894,549**		**582**	**1,695,538**		**3,937**	**5,590,087**	
Elderly living alone; % of persons aged 75 and older living alone										
Least 75+ living alone (Q1)	0.0% -24.4%	843	1,003,543	26%	37	110,246	7%	880	1,113,789	20%
Quintile 2	24.5% -30.6%	681	897,794	23%	83	222,320	13%	764	1,120,114	20%
Quintile 3	30.7% -35.7%	589	796,720	20%	108	322,087	19%	697	1,118,807	20%
Quintile 4	35.8% -42.5%	663	745,285	19%	134	373,596	22%	797	1,118,882	20%
Most 75+ living alone (Q5)	42.6% -100.0%	579	451,207	12%	220	667,288	39%	799	1,118,495	20%
Total		**3,355**	**3,894,549**		**582**	**1,695,538**		**3,937**	**5,590,087**	

** Annual average of combined 1995-2007 population.*

CORRELATION BETWEEN SOCIO-ECONOMIC MEASURES AND POPULATION DENSITY

The three socio-economic measures and population density had varying degrees of correlation. However while the correlation coefficients between several of the measures were statistically significant, none represented a high level of correlation. The highest correlation was a negative association between education and population density (-0.365) (Table 2.6).

Table 2.6 Correlation coefficients (Spearman's rank) for ward/ED characteristics

	% of economically active persons aged 16-74 who were unemployed	% of persons aged 16-74 without a university degree (or academic equivalent)	% of persons aged 75 and over living alone
Population density (persons per hectare)	0.231	-0.365	0.112
% of economically active persons aged 16-74 who were unemployed		0.197	0.147
% of persons aged 16-74 without a university degree (or academic equivalent)			0.008

GEOGRAPHIC DISTRIBUTION OF SOCIO-ECONOMIC MEASURES AND POPULATION DENSITY

Map 2.3 shows Ireland divided into approximate population density tertiles (<1 person/hectare, 1-15 persons/hectare and >15 persons/hectare). As expected, only geographic areas at the centre of large towns and cities, such as Belfast and Dublin, fell into the highest tertile. The majority of wards/EDs in Ireland had a population density of less than 1 person/hectare.

Map 2.4 shows the percentage unemployed in each ED/ward by quintiles. Areas of highest unemployment were found in north and west Belfast, north-west Ireland (including Donegal, Derry and Strabane), the west of Ireland (including Mayo) and parts of Newry & Mourne and Louth.

Low levels of tertiary-level education (as illustrated in Map 2.5) were found in rural parts of RoI, north and west Belfast, north-east Dublin and south-west Dublin. High levels of tertiary education were found in south Belfast and surrounding areas, central and southern Dublin and surrounding areas, and other urban areas and their environs in RoI, such as parts of Cork, Galway and Limerick.

Areas with high proportions of elderly persons (aged 75 and over) living alone were fairly randomly spread across Ireland, as seen in Map 2.6. The proportion was relatively high in Dublin and Belfast city centres, but low in the surrounding areas.

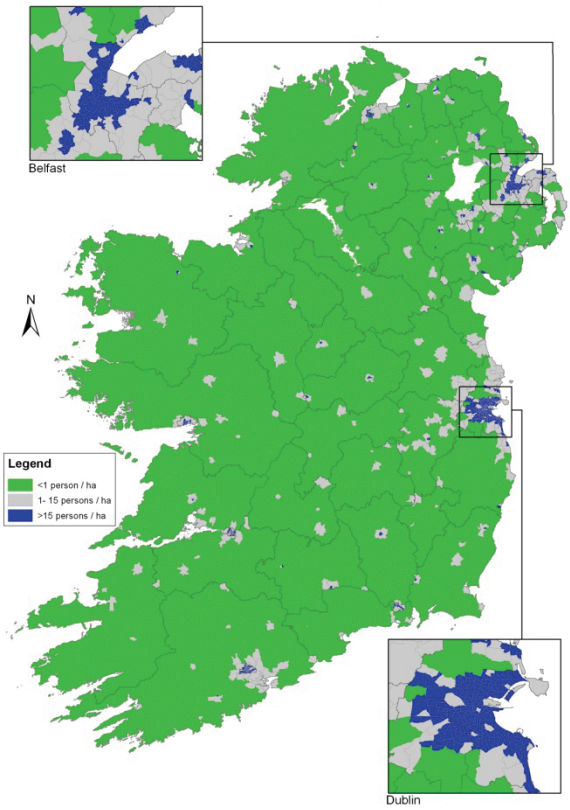
2.2.5 INTERNATIONAL CANCER INCIDENCE DATA

Estimates of cancer incidence rates in 19 (mainly European) countries, of a level of economic development comparable to Ireland, were taken from the GLOBOCAN 2008 dataset (Ferlay et al., 2008). Data for NI is included in that for the UK in this dataset, so the incidence rates shown for both NI and RoI are based on 2005-2007 data from the respective cancer registry. For this reason, the ranking of RoI and NI relative to each other shown in these figures is not always the same as that shown in the summary section, or in the sections for each cancer titled “Small geographic area characteristics and cancer risk”. It should also be noted that, although countries are shown as ranked in descending order of incidence rates, the differences in rates between countries were often quite small and may not be statistically significant.

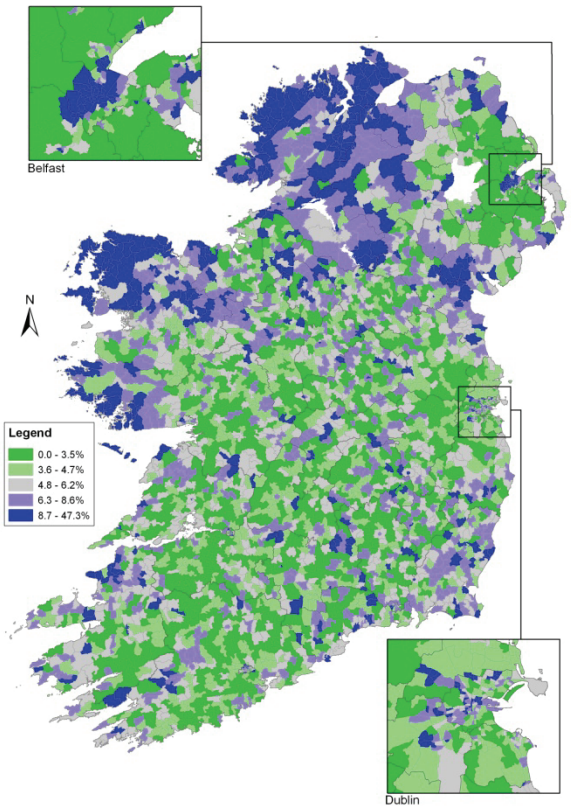
For some cancers, the definition of cancer site used in GLOBOCAN differed slightly from that used in this atlas. In these cases international comparisons between Ireland and other countries are based upon the GLOBOCAN definition. The exception to this was non-Hodgkin’s lymphoma, where the ICD code C96 was included in the GLOBOCAN definition, but omitted from the NI and RoI figures. A footnote to the comparison graphs is provided to indicate where such differences occur.

Maps 2.3-2.6 Ward/ED characteristics

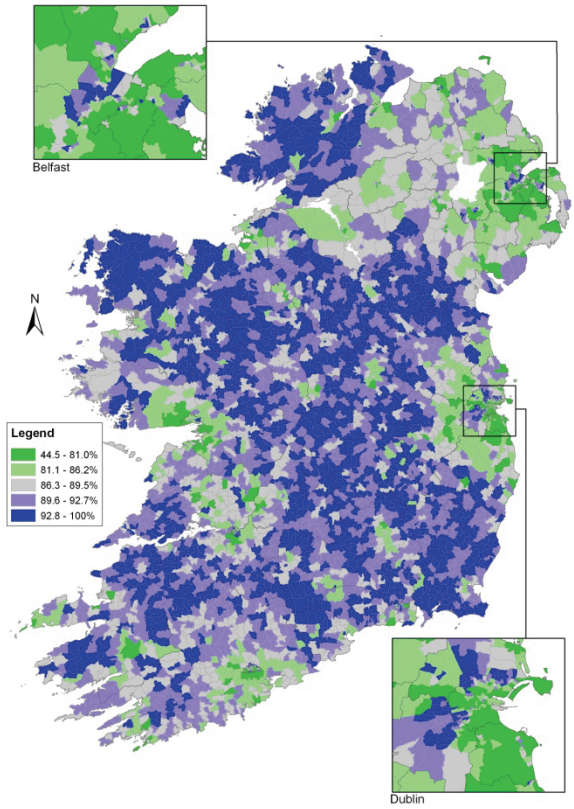
Map 2.3 Population density (persons/hectare)



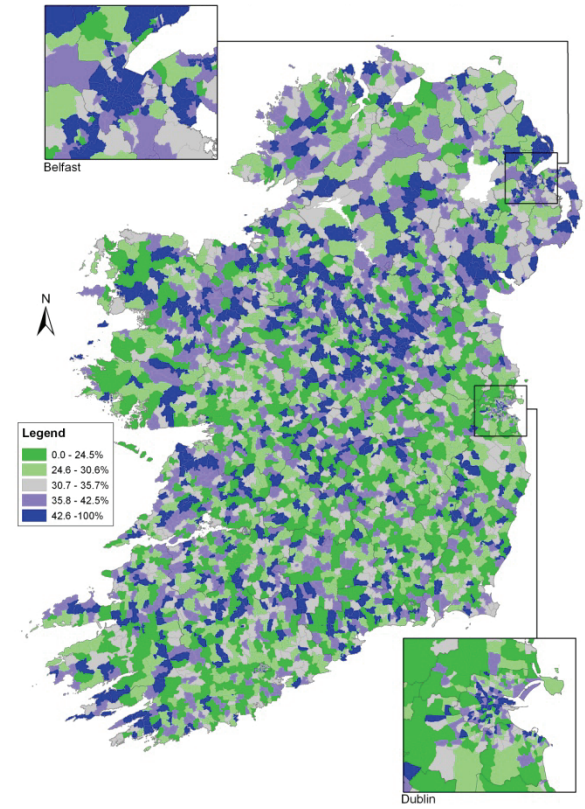
Map 2.4 % unemployed



Map 2.5 % of population without a degree



Map 2.6 % aged 75 and over living alone



2.3 STATISTICAL METHODS

2.3.1 STANDARDISED INCIDENCE RATIO

In comparing cancer incidence between areas or over time, three important factors must be considered—the number of people at risk, their sex and their age. In this report, cancer incidence for men and women was considered separately, which deals with possible differences between sexes. The reason for correcting for the number of people at risk is obvious; the number of cases is divided by the number of people resident in the area during a specified period (as reported by the census) to produce an incidence rate.

Since the risk of developing cancer doubles with every eight or nine years of life, an area with an older population would be expected, all else being equal, to have more incident cancer cases than an area with a younger population. There are several different approaches available to adjust for differences in age; this atlas has used indirect standardization, which is the most appropriate method for small area comparisons, as it provides more stable rates than other standardization techniques, and works even if there is no population-at-risk in some age groups within the area (Estève et al., 1994). For each small area i, the national incidence rates for each age group j were applied to the population counts (N) in each age group, to calculate the total expected number of cancers (E) in the area. This can be compared to the number actually observed (O) in the area, in the form of an observed to expected ratio, or percentage. This is called the standardised incidence ratio, abbreviated to SIR. The SIR for any cancer for either men or women for Ireland as a whole is, by definition, 1 (or 100%), where for any small area (ED or ward) i:

$$SIR_i = \frac{O_i}{E_i} \quad \text{where } E_i = \sum_{\text{all age groups } j} N_{ij} \frac{O_j}{N_j}$$

2.3.2 SPATIAL ANALYSIS AND SMOOTHING

There are several types of geographical analysis of disease incidence:

- disease mapping, which aims to provide an estimate of the disease rate in each small area which is as close as possible to the true value;
- cluster studies, which specifically search for “clusters”—areas or groups of areas where risk is significantly higher than in the rest of the population;
- point source studies, which investigate disease risk around a "point source" of possible risk which has been defined a priori (e.g. an industrial site).

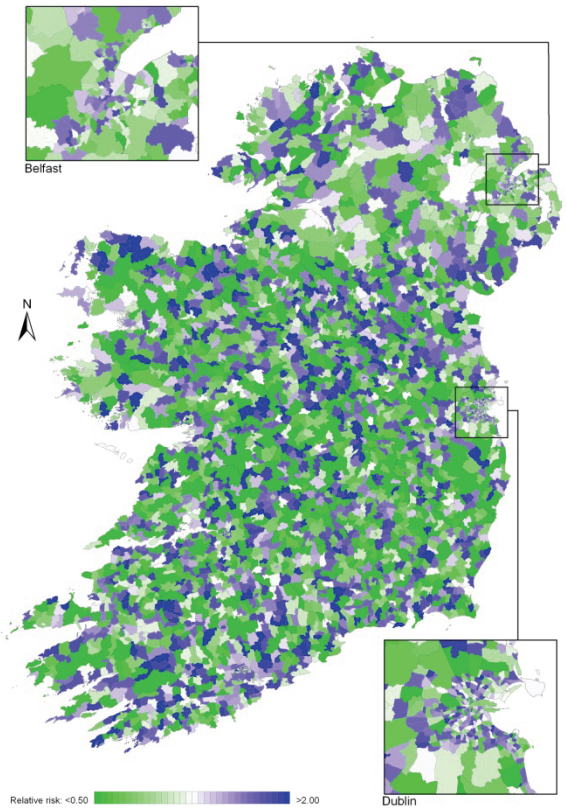
Because the primary aim was to estimate risks precisely in each small area (ED or ward), disease mapping methodology was used.

Incidence rates, whether crude or standardised, are subject to high variability due to the small number of cases occurring in each small area, and the often small population-at-risk. In many instances, areas with small populations can appear to have a particularly high or low risk, purely by chance. The average population of an ED or ward in Ireland overall was about 1,420, but some were considerably smaller. One of the commonest cancers, colorectal cancer, had an incidence rate of 0.5 cases per 1,000 persons per year, so even over the 13-year period examined here, only about 9 cases would be expected in an average ED or ward, and most cancers analysed in this report have considerably lower incidence rates than this. With such small numbers, random variation is the major factor in the variation of incidence rates between EDs or wards, and this “noise” tends to obscure any other patterns. Therefore, simply mapping the SIRs for each ED or ward can be seriously misleading, as the SIRs tend to be more extreme in areas where the population is sparse. These areas are often the largest in area and can dominate a map visually. This is illustrated for colorectal cancer in men in Map 2.7.

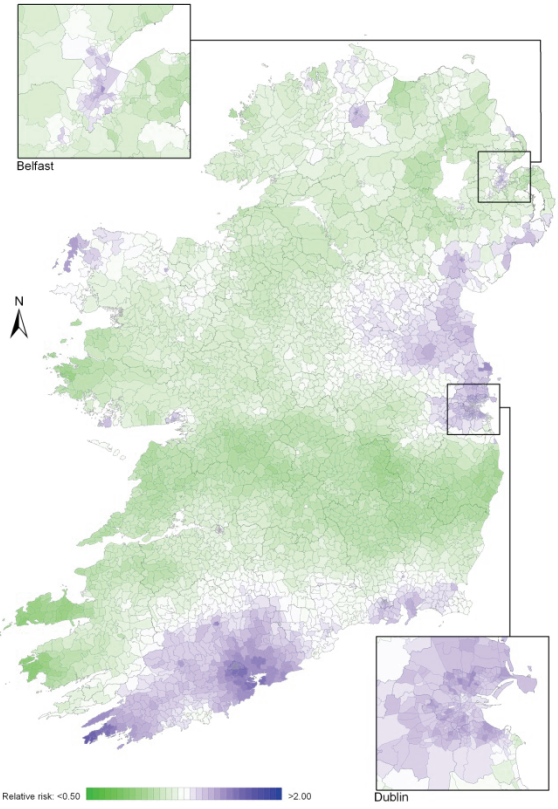
The way of dealing with this problem involves "smoothing" the estimates of disease risk (Elliott et al., 1996). Smoothing removes the noise (i.e. it smoothes out the random variation) and shows more clearly the geographical pattern of the true underlying distribution of cancer rates—or the relative risks (RR). The effect of

smoothing is illustrated in Map 2.8, which shows smoothed RRs for male colorectal cancer, compared with the unsmoothed SIRs in Map 2.7.

Map 2.7 Colorectal cancer, crude standardised incidence ratios: males, 1995-2007



Map 2.8 Colorectal cancer, smoothed relative risks: males, 1995-2007



The principle of spatial smoothing is straightforward. If we assume that the risk of cancer does not vary much between areas which are close to each other, then differences between EDs or wards are more likely to be due to random variation than to real differences in risk. The smaller the population of the area, the larger will be the element of random variation and the crude SIR will be quite an unreliable indicator of real risk. Smoothing the SIR for an ED or ward allows us to strengthen the estimate for the ED or ward by "borrowing strength" from adjacent areas (local smoothing) and/or from the overall/national map (global smoothing) in order to increase the stability of the estimated RR. Therefore, smoothing adjusts risk estimates based on small numbers towards a local mean—based on the rates in the neighbouring areas—and also towards the national value.

Many methods have been proposed for smoothing disease rates (Elliott et al., 1996; Best et al., 2005). We have chosen to use a Bayesian approach (Best et al., 2005). The main advantage of Bayesian techniques is that they work well in situations of limited information and high uncertainty. They are better at accurately depicting the geographical pattern in risk than other techniques, such as non-hierarchical approaches, which are more likely to be visually misleading (Pascutto et al., 2000).

The SIRs were smoothed by estimating relative risks using conditional autoregressive models (CAR) (Clayton and Kaldor, 1987) based on a spatial Poisson model with two random effects, as follows:

$$O_i \sim Poisson\ (E_i \vartheta_i)$$

$$log\ (\vartheta_i) = \alpha + h_i + \theta_i$$

where

O_i was the observed number of cancer cases in area *i;*

E_i was the expected number based on age-adjusted national incidence rates in area *i*;

ϑ_i was the estimated relative risk in area *i*;

α was the intercept;
h_i was a random effect which models the unstructured heterogeneity; and
θ_i was a spatially structured random effect (which is given a CAR prior distribution).

Use of CAR models is widespread in disease mapping and this particular model is considered to be appropriate in most situations (Lawson et al., 2000; Best et al., 2005). The suitability of the specific model above for Ireland was evaluated by comparing it with several alternative models which included covariates for population density and/or country. However, it was decided to use the basic model in this atlas as, while the alternative models were successful in detecting covariate effects, it was not clear what the covariates were actually markers for. Any effects due to socio-economic factors, for example, would be identified by means of the negative binomial regression analysis (section 2.3.3).

Other disease mapping methods (e.g. kernel smoothers, mixture models) seem to give poorer results than CAR (Lawson et al., 2000). Although risk estimates can be somewhat underestimated, CAR models have a high specificity (Richardson et al., 2004), and this conservative approach means that high or low estimates are more likely to be real. However, as with any smoothing method, it is possible that areas of genuinely high risk may be missed by smoothing with neighbouring areas. The method also assumes that risk varies smoothly at the scale studied, an assumption which may not be justified if risk factors vary considerably at a purely local level.

Models were fitted using Markov Chain Monte Carlo (MCMC) algorithms with WinBUGS software (Lunn et al., 2000). Estimates were checked to ensure convergence had been reached. A burn-in of 150,000 iterations was performed and the posterior distributions were derived using one in three iterations from the subsequent 10,000 iterations of 2 chains.

Ireland has a number of off-shore islands which form EDs but which have no neighbours (i.e. adjacent areas). Smoothing is based on a shared boundary between EDs, and the absence of such a boundary means that the risk for islands cannot be smoothed in the same way as that for mainland EDs. A similar situation arises with a number of headlands and small peninsulas, which share a boundary with only one other ED. It is common for such EDs or wards to appear as “hotspots” on smoothed maps. To minimise this problem, we created artificial “neighbours” for islands and those headlands which had only one neighbour, by assigning the nearest mainland EDs or wards as “additional neighbours”, so that each island and headland had a minimum of two neighbours (Appendix table A2.5). The “additional neighbours” were given a weighting half that of true neighbours in the smoothing algorithm.

Relative risks (RR) were mapped for each cancer site individually using ArcMap 9.3. For those cancers which affect both sexes, maps are included for both sexes combined and for men and women separately. County and district council boundaries are shown faintly on the maps to help the reader with geographical orientation; a map of these is on page 4 (Map 2.1). To aid orientation, a map is also provided at the same scale, showing the same boundaries, as well as some towns and cities on the island (Map 2.2). To facilitate comparisons between cancer sites, each map is shown using the same colour ramp, which ranges from dark green for an estimated RR less than 0.50 to dark blue for a RR higher than 2.00 (i.e. the same colour represents the same value of RR on each map). The grid from 0.50-1.00 was based on the assumption of normality of the estimated relative risks so that approximately equal numbers would fall into each interval. The grid from 1.00-2.00 was chosen as the reciprocal of the 0.50-1.00 intervals (e.g. the reciprocal of 0.50-0.55 is 1.82-2.00) as this was considered appropriate for ratios (relative risks). This scale is different from that used in the RoI atlas (Carsin et al., 2009) and so the maps are not directly comparable.

Appendix table A3.1 contains summary information from the mapping of each cancer site, including average numbers of cases per ED and ward, and ranges of SIRs and smoothed RRs.

2.3.3 REGRESSION ANALYSIS: WARD/ED CHARACTERISTICS AND CANCER INCIDENCE

A count of the number of cases of cancer by type and sex was available for each ward/ED. Relating these counts to the ward/ED characteristics is traditionally done by modelling the count data using Poisson regression. However a key assumption behind this approach is that the mean and variance of the counts being modelled are the same. Deriving the mean number of cancer cases diagnosed in each small geographic area, and the variance between areas in these counts, illustrates that this assumption is not valid and that the data is over-dispersed; that is, the variance is greater than the mean (Table 2.7) (Breslow, 1984).

Table 2.7 Mean and variance in the number of cancer cases diagnosed in each ward/ED: 1995-2007

cancer	males		females	
	mean	variance	mean	variance
non-melanoma skin cancer	14.2	347.5	12.5	334.7
breast	–	–	9.8	186.5
colorectal	5.4	48.9	4.3	35.0
lung	5.3	58.7	3.3	30.7
prostate	8.4	108.3	–	–
non-Hodgkin's lymphoma	1.3	3.7	1.2	3.4
stomach	1.5	5.0	0.9	2.7
melanoma of the skin	0.9	2.5	1.4	5.3
bladder	1.6	5.4	0.6	1.4
head and neck	1.4	5.3	0.6	1.2
leukaemia	1.2	2.9	0.8	1.9
pancreas	0.9	1.9	0.9	2.3
kidney	1.0	2.7	0.6	1.3
oesophagus	1.0	2.5	0.6	1.3
ovary	–	–	1.6	5.7
brain and other central nervous system	0.8	1.6	0.6	1.1
cervix uteri	–	–	1.0	3.1
corpus uteri	–	–	1.3	4.1

Although a great deal of this variance may be explained by the differing population sizes of each geographic area, which is adjusted for in a Poisson regression model, we decided to use a modification of Poisson regression, known as negative binomial regression, to adjust more fully for the over-dispersion. This model produces a relative risk (RR) for each categorical variable included in the model, relative to a baseline value. For example, if RoI is taken as the baseline (by definition, RR=1) in a variable indicating which country the geographic area is in, then if NI has a relative risk greater than 1, this means that the incidence of cancer is higher in NI than RoI; conversely a relative risk lower than 1 means that incidence is lower in NI than RoI. Five small area characteristics were examined for a relationship to cancer incidence using this approach—country, population density tertile, and quintiles of unemployment, third-level education and elderly living alone (see section 2.2.4.3).

It has already been noted (section 2.2.4.3) that the variables we are studying are not completely independent of each other. Therefore, if we see a relationship between cancer risk and a specific variable (for instance level of unemployment), part of this relationship might be due to another factor, such as the average age of the population, which would influence both cancer rates and unemployment levels. For this reason, measures of the effect of each variable must be adjusted for the effects of the others (see section 2.3.1). The most important adjustment is for age, as cancer risk rises rapidly with age. Two comparisons were made between NI and RoI, one of which was adjusted for age alone, and the other for age, population density, unemployment, education and percentage of elderly living alone. All other relative risks reported were adjusted for the effects of all the other variables. Thus, risk estimates are reported for:

- country adjusted by age only;

METHODS

- country, adjusted by age, population density, unemployment, education and elderly living alone;
- population density, adjusted by age, country, unemployment, education and elderly living alone;
- unemployment, adjusted by age, country, population density, education and elderly living alone;
- education, adjusted by age, country, population density, unemployment and elderly living alone; and
- elderly living alone, adjusted by age, country, population density, unemployment and education.

The risk estimates with 95% confidence intervals and tests of statistical significance are given in full for each site in Appendix 1. Summary figures are presented in each chapter.

2.3.4 SUMMARY MEASURES

A series of summary measures was computed for each cancer site. The incidence of each cancer is expressed in terms of the average number of new cases each year between 1995 and 2007, and as a percentage of all new cancer cases, both including and excluding non-melanoma skin cancer.

TIME TRENDS

Estimated annual percentage rate of change in the number of cases was calculated over the period 1995-2007 (13 years) by taking the 12th root of the total percentage growth rate (12 years of growth).

$$\text{Annual \% rate of change} = 100 \times \left(\left(\frac{\text{No.cases in 2007}}{\text{No.cases in 1995}} \right)^{\left(\frac{1}{12}\right)} - 1 \right)$$

CUMULATIVE RISK

Cumulative risk to age 74 ($\hat{R}_{74}$) is the risk of developing a specified cancer or cancers up to and including age 74, in the absence of competing risks (Estève et al, 1994). This was calculated as follows:

$$\hat{R}_{74} = 1 - e^{-t_{0,74}}$$

where, if x is one of 15 five-year age groups from 0 to 74:

$$t_{0,74} = 5 \sum_{x=1}^{15} t_x$$

t_x=age-specific incidence rate

The cumulative risk is given as a percentage and also as a ratio (e.g. a cumulative risk of 4% is expressed as 1 in 25).

PREVALENCE

15-year prevalence was estimated as the total number of individuals diagnosed between 1/1/1994 and 31/12/2008 who were still alive on 31/12/2008. Numbers are given for those who were aged under 65 years on 31/12/2008, and for those who were aged 65 years or older on that date.