Appendix 4: Registry data set

The Registry dataset has been continually extended since 1994, in responses to the changing needs of our data users. With the introduction of a new computer system late this year, we have added a number of new data items. Most of the data we intend to collect is listed below. We would be interested to hear from any data users with comments on the usefulness or definitions of the items below, or with suggestions for new data items.

Patient data

This table contains patient specific data. All identifiable information is stored here.

VARIABLE NAME	DESCRIPTION
registration number	Uniquely identifies a registration
year	The year the registration was made
first name	First name of patient being registered
second name	Second name of patient
surname	Surname of patient
soundex	'Sounds like surname' code. Automatically generated by system
maiden name	Maiden name of patient
date of birth	Date of birth
year of birth	In the absence of an exact date of birth, the year of birth may be known
sex	Male / female
GMS number	GMS number
PPS number	PRSI number
occupation code	Standard occupational classification code
occupation status	Retired / student / self-employed / employed / housewife / religious / other / unknown
whose occupation	Person whose occupation is described: own / husband / wife / partner / father / mother / parent / unknown
marital status	Single / married / divorced / widowed / separated / unknown/other
smoker	Yes / no / ex-smoker / unknown
dead	Yes / no
date of death	May be entered by TRO or updated automatically by system when processing death certificates
year of death	If exact date of death is not known, this data item stores the year of death
cause of death	ICD-9 code
death certificate number	Uniquely identifies a death certificate
patient address	Uniquely identifies this address for this patient
main address	This is set to 1 if the main address, 0 otherwise
house number	House number
house name	House name
address 1	First line of address
address 2	Second line of address
address 3	Third line of address
address 4	Fourth line of address
county id	County code, which is validated against the county lookup table
DED	District electoral division of residence
screening status	Screening unspecified / organised / opportunistic

Tumour data

This table contains all information specific to each tumour. A patient may have multiple tumour records.

VARIABLE	DESCRIPTION
registration number	Patient registration number, which is linked to the patient table
tumour id	The unique identifier for this tumour which is system generated
progression	For tumours that have progressed this indicates the progression level. For analysis the tumour with the highest progression number will be included, previous records of the tumour will not
practice	The code of the general practice that the patient attends
GP id	The GP code
date of incidence	Date of diagnosis
age at incidence	The age of the patient at diagnosis date
topography code	Site of primary tumour; ICD-o-2 code
topography id	Uniquely defines the description of the ICD-o-2 topography code selected
morphology code	Histological type of tumour; ICD-o-2 code
method of diagnosis	Histology of primary/histology of other site/cytology/marrow/blood film/tumour marker/clinical-visualisation/clinical-no visualisation/clinical-unknown/radiology/post mortem/other/unknown
side	Right/left/both/midline/unknown
grade	Histological grade
method of presentation	Symptoms /incidental/screening-unspecified/screening- organised/screening-opportunistic/autopsy/unknown
histology lab number	Histology lab number
histology date	Date of histology
pathologist id	Name of pathologist table
pathology lab	Lab providing report
clinical stage T	Clinical TNM stage
clinical stage N	Clinical TNM stage
clinical stage M	Clinical TNM stage
pathological stage T	Pathological TNM stage
pathological stage N	Pathological TNM stage
pathological stage M	Pathological TNM stage
TNM-summary	Calculated from the entries given on the TNM stage data-items above
confidence T	Most reliable basis of T stage
confidence N	Most reliable basis of N stage
confidence M	Most reliable basis of M stage
residual	Residual disease. Valid values are none/microscopic/macroscopic/not applicable/cannot be assessed
extent of disease	Local/regional/distant
occurrence	The number of occurrences of this tumour. Specific to skin (C44) cancers
tumour size	Size of the tumour (mm)
tumour marker	Marker type and value
GP referral date	Date on which GP requested a hospital appointment
first appointment date	Date of first hospital appointment/attendance
first recurrence date	Date of first recurrence (if any)
source of notification	Source of information: valid values are pathology/death cert/GP/radiotherapy/other outpatient/other inpatient/central sources/HIPE/unknown
date of notification	Date this source was accessed
hospital id	Hospital of diagnosis
metastasis id	Unique metastasis id for this tumour
topography code	Valid ICD-o-2 code for metastasis
date of metastasis	Date of metastasis

Management

This table describes each significant recorded contact between the patient and the medical services. Initially only primary treatment episodes are being recorded.

VARIABLE	DESCRIPTION
admission type	Elective inpatient/emergency inpatient/outpatient visit/day admission
type of procedure	Primary treatment/recurrence/metastasis/palliative/other
consultant id	Consultant caring for the patient
MRN	Medical record number
topography code	Site of procedure
treatment code	ICD-9-CM code
date field	Date of treatment