EUROCARE-5

PROTOCOL FOR UPDATING POPULATION-BASED

CANCER SURVIVAL IN EUROPE

10 March 2010

INDEX

1. INTRODUCTION

EUROCARE is a collaborative research programme set up in 1989 and currently involving population-based cancer registries in 23 European countries. Most of these countries are Member States of the European Union. EUROCARE monitors trends in the survival of cancer patients in participating countries. It also evaluates the outcome of the care of cancer patients (on large random samples).

The first EUROCARE study included patients diagnosed during 1978-85 and followed up to 1989. It was subsequently extended to include patients diagnosed during 1985-89 (EUROCARE-2), 1990-94 (EUROCARE-3) and 1995-2002 (EUROCARE-4)[1,2,3,4]. This protocol refers to EUROCARE-5, which will include European cancer patients diagnosed up to 2007.

The EUROCARE Co-ordinating Centre is at the Istituto Nazionale Tumori (INT) in Milan (Italy). Data checks and basic analyses are performed at the EUROCARE Data Analysis Centre at the Istituto Superiore di Sanità (ISS) in Rome (Italy), where the data are stored.

1.1 Aims of the project

This protocol incorporates decisions taken by the EUROCARE Steering Committee in Utrecht on 12 December 2008, and the meeting of the EUROCARE Working Group in Genova, 10-12 March 2009.

Major changes in European cancer patient survival in the decade following 2000 are expected as a result of the introduction of new targeted treatments and new diagnostic tools. In several European countries, changes in health care systems and implementation of national cancer plans and screening programmes are also likely to have influenced patient survival.

EUROCARE-5 will update and expand the EUROCARE database in order to:

- continue monitoring population-based cancer survival in Europe;
- include data from additional cancer registries and countries if possible;
- continue monitoring variations in cancer survival by country, region, age, time and sex;
- extend the follow-up of cancer patients included in previous analyses to study both long-term survival and temporal trends in survival;
- use the available information on stage at diagnosis and treatment to interpret survival variations and time trends;
- use the period survival method to provide short-term predictions of survival based on the most recent follow-up data;
- update the estimates of cancer prevalence;
- estimate the proportion of cancer patients who are cured of their disease;
- estimate the number and proportion of avoidable deaths;
- compare relative and cause-specific survival by using available information on the underlying cause of death;
- reduce the delay in reporting population-based survival by improving data submission procedures;
- make datasets available to participating EUROCARE centres for further analyses: the Co-ordinating Centre and the Data Analysis Group will assist researchers interested in such studies.

1.2 Resources

Resources (personnel, computers and facilities) have been allocated by the INT in Milan and the ISS in Rome for co-ordination of the EUROCARE project and analysis of the data. A dedicated server is available at ISS for data storage and analysis. Limited funds for meetings and travel are available at the two Institutes, although a specific budget cannot be stated in advance. Resources for specific, additional initiatives will be sought as required.

1.3 Confidentiality, security and ethical approval

As in the previous EUROCARE studies, cancer data will be stored individually, but anonymously. Data will be stored in a dedicated server that is not connected to the web, and according to the standard requirements for data security at ISS (Rome). Data handling conforms with the confidentiality guidelines published by the International Association of Cancer Registries (www.iarc.fr). The EUROCARE protocol received institutional approval as part of the CONCORD project from the Scientific Ethical Committee of the Istituto Superiore di Sanità in 2002.

1.4 Outcome and publication policy

The main output from EUROCARE-5 will be the publication of survival estimates for cancer patients diagnosed during 2000-2007. By May 2010, most registries will be able to provide data on patients diagnosed up to 31 December 2007, with follow-up to 31 December 2008. During 2010, these data will be checked and analysed by the methods developed previously. First reports with the summary results will be ready by early 2011. More specific analyses and publications will follow soon afterwards; a schedule will be circulated after most or all data sets have been received.

The EUROCARE publication policy incorporates guidelines that have been approved by all participants, and was last updated in September 2004. It is available on the EUROCARE web-site (www.eurocare.it).

2. CRITERIA FOR INCLUSION

All invasive, primary, malignant neoplasms, except non-melanoma skin cancer, are eligible for inclusion in the EUROCARE database.

In addition, ***in situ*** cancers of **breast, cervix, colon-rectum and skin (melanoma)** are also eligible for inclusion. The frequency of *in situ* malignancies at these sites is an indicator of the intensity of screening and early diagnostic activity, which are both increasing in many countries, and previous analyses suggest that this can also improve the interpretation of international differences in survival from invasive cancers.

Further, in accordance with the procedures used by many cancer registries, **benign** tumours of the **central nervous system** and **benign and *in situ*** tumours of the **urinary bladder** are also eligible for inclusion.

The tumours defined as above will be collectively described as 'index tumours'. Data for other tumours should not be supplied.

Data for all index tumours will be collected, including death-certificate-only (DCO) cases, those discovered at autopsy, and those lost to follow-up. Anatomic site and tumour morphology and behaviour must be coded according to the **third edition of the International Classification of Diseases for Oncology (ICD-O-3)**, published in 2000. Both microscopically verified and non-verified cases must be included.

Multiple primary tumours occurring in the same person must also be included. Unlike previous EUROCARE studies, EUROCARE-5 will include second and higher-order (third, etc.) primary malignancies in the main analyses of cancer survival. Therefore each tumour record will require both a tumour identification code and a patient identification code.

At the last EUROCARE Working Group meeting it was agreed that expanding the availability of clinical variables, particularly those relating to **stage at diagnosis,** should be a priority for EUROCARE-5. To reduce the practical difficulties of collecting this information, it was also agreed to restrict the collection of information on stage to selected solid tumours: stomach, colon-rectum, skin melanoma, breast, cervix uteri, corpus uteri, ovary, prostate, testis, kidney, bladder and thyroid.

3. PATIENT DATA

Data should be supplied on all index tumours (as defined in Section 2) diagnosed up to **31 December 2007** in patients who were resident in the territory covered by the registry, with their vital status updated to **31 December 2008**.

Registries are requested to send updated records for **all index tumours**, including those that were included in the EUROCARE-4 study. For tumour records that were submitted for EUROCARE-4, we recommend retaining the **same identification codes**, in order to simplify data quality control procedures.

Survival up to 5 years will be estimated with the cohort approach for patients diagnosed during **2000-2004**. Short-term predictions of survival up to 5 years will be made with the period approach for patients diagnosed during **2005-2007.**

For each patient the following variables are required:

3.1 IARC check flag (one-digit variable) **Compulsory**

1 = Record checked according to the IARC CHECK program
2 = Record not checked

This field will be used to avoid unnecessary requests to check records that have already been checked and verified by the registry. However, it is ***not required*** that this program be run by the cancer registry before submission of its data.

3.2 Sex (one-digit variable) **Compulsory**

Records with a missing or invalid code for sex will be counted and excluded. Records will be checked for sex-site incompatibility.

1 = Male
2 = Female
9 = Sex is ambiguous or unknown

3.3 Date of birth **Compulsory**
(two-digit variable for month, four-digit variable for year)

Month = 01-12
Year = **1900-2010 (plausible range)**

3.4 Date of diagnosis **Compulsory**
(two-digit variables for day and month, four-digit variable for year)

The date of diagnosis should be consistent with that used for computing incidence during 2000-2007. Each registry should describe how the index date was defined for this purpose.

Day = 01-31
Month = 01-12
Year = **1978-2010 (possible range)**

The day and month of diagnosis are essential for accurate survival analyses. The day of diagnosis is required to enable analysis of survival in the earliest follow-up intervals (first three months after

diagnosis) and to study outcome indicators such as early death. If the day is not available, that field must be left empty.

3.5 Date of tumour registration **Optional**
(two-digit variables for day and month, four-digit variable for year)

This information is optional. It is the date when the tumour was added to the registry database, *not* the date of diagnosis. It is requested only from registries in which it is routinely collected. It will be used to study the frequency of late registration and the survival of patients registered late, in order to estimate any bias in overall survival that could be attributable to reporting delay.

Day = 01-31
Month = 01-12
Year = **1978-2010 (possible range)**

3.6 Date of death or of last known vital status **Compulsory**
(two-digit variables for day and month, four-digit variable for year)

Participating centres will follow up patients using their usual procedures. This field must be completed with the date of death or of last known vital status. Vital status must be ascertained at **31 December 2008** or later for the **entire cohort of patients**. For the small percentage of patients lost to follow-up, in whom vital status cannot be verified at follow-up closing date, the last date on which the case was known to be alive must be reported. If the day is not available, that field must be left empty.

Day = 01-31
Month = 01-12
Year = **1978-2010 (possible range)**

3.7 Vital status (one-digit variable) **Compulsory**

1 = Alive at the date of last known vital status
2 = Dead (the date of death is given in field 3.6)
3 = Lost to follow-up. For these cases, the last date at which they were known to be alive should be given in field 3.6. If no date is known, report the date of diagnosis
4 = Death-certificate-only (DCO) case. For these cases, the date of diagnosis is the same as the date of death
5 = Case detected only at autopsy. For these cases, the date of diagnosis is also the same as the date of death

3.8 Primary tumour site (three-digit variable) **Compulsory**

The anatomic site of the primary tumour should be coded to the **third revision of the International Classification of Diseases for Oncology (ICD-O-3)**. Coding to ICD-9 or ICD-10 classifications **should not be used.**

Only the numeric part of the ICD-O-3 topography code must be reported (do not include the “C”). Do not include the decimal point. The valid range for this field is:

Site = 000-809

3.9 Microscopic confirmation of diagnosis (one-digit variable) **Compulsory**

1 = Histologically confirmed
2 = Cytologically confirmed
3 = Microscopically confirmed, but not known whether by histology or cytology
4 = No microscopic confirmation
9 = Unknown

3.10 Morphology (four-digit variable) and behaviour (one-digit variable) Compulsory

The morphology and behaviour of the primary tumour should be coded to the **third revision of the International Classification of Diseases for Oncology (ICD-O-3)**. Coding to ICD-O-1 or ICD-O-2 **should not be used.**

Morphology code in the range 8000-9989
Behaviour code:

0 = Benign
1 = Uncertain whether benign or malignant
2 = Carcinoma in situ
3 = Malignant, primary site

Registries that use behaviour codes 6 and 9 are requested to explain how they use these codes in an accompanying letter.

3.11 Summary extent of disease at diagnosis (one-digit variable) Compulsory

1 = Tumour is confined to the site of origin
2 = Tumour has spread to immediately adjacent tissues and/or regional lymph-nodes
3 = Tumour has spread to distant organs
4 = Tumour is not confined to the site of origin but not specified whether code 2 or 3 applies
5 = No distant metastasis but not specified whether code 1 or 2 applies
9 = Unknown summary extent of disease

3.12 Patient identification code (ten-character field) Compulsory

The anonymous patient identification code used by the cancer registry to refer to each person registered with a cancer in its own database must be included in each tumour record, to facilitate quality control. It will be used to enable resolution of problems identified by the quality control procedures, in discussion with the registry.

This code can be any unique string of characters, but not the person's name or any national identity number. The same code must be included in any other tumour record(s) supplied for the same person, to enable distinction between first and second (or later) tumours. For patients whose data were submitted to EUROCARE-4, the patient identification code should be the same as in that submission, to avoid ambiguity in quality control and in tracing multiple tumours.

Up to ten numeric or alphanumeric characters are allowed.

3.13 Tumour identification code (ten-digit field) Optional

The unique tumour identification code used by the cancer registry to refer to each tumour registered in its database must be included. Taken together with the person code (field 3.12), this code will enable multiple primary tumours in the same person to be separated in the analyses. The code will also be used to enable resolution of problems identified by the quality control procedures, in discussion with the registry. Up to ten digits are allowed.

For tumour records that were submitted for EUROCARE-4, the tumour identification code should be the same as in that submission, to avoid ambiguity in quality control and in tracing multiple tumours. [Note: a few registries only submitted the tumour identification code, but no person identification code (field 3.11) for EUROCARE-4, and they will be contacted separately to resolve this issue].

3.14 Multiple tumour code (two-digit variable) **Compulsory**

Some patients are registered with more than one tumour. This field specifies the rank (first, second, third, etc.) of the tumour for a given patient. It will enable analyses to be performed with and without the inclusion of second, third (etc.) tumours. The main survival analyses in EUROCARE-5 will include ***all*** **index cancers diagnosed during the period 2000-2007, not just the *first* index cancer.** Sensitivity analyses will be done to assess the impact of this change from previous practice.

Index cancers **include** benign tumours of the **central nervous system** (CNS) and benign and *in situ* tumours of the **urinary bladder**. These tumours should be included as part of multiple primary malignancy sequences.

Index cancers **do not include** any other benign or *in situ* tumours, or non-melanoma skin cancer. Those tumours should not be included in the enumeration of multiple tumours.

00 = single tumour (person with only one primary malignancy recorded)
01 = first tumour (first primary malignancy in person with two or more primary malignancies)
02 = second tumour (second primary in person with two or more primary malignancies)
03 = third tumour (third primary in person with three or more primary malignancies)
[etc.]

3.15 Stage at diagnosis **Optional**

Stage at diagnosis is increasingly important for international survival comparisons, particularly for comparison of survival between Europe and the USA. In many registries, information of satisfactory quality on stage at diagnosis is now available for most of the cancers included in EUROCARE-5. We will perform specific analyses of survival in relation to stage at diagnosis.

When stage data coded according to the TNM classification are available, they should be reported. When TNM staging information is not complete, it is proposed to record the condensed TNM as recommended by the ENCR Working Group on extent of disease (http://www.encr.fr/). If neither TNM nor the condensed TNM are available, the only information on stage to be given is summary extent of disease (field 3.11). The pathological stage (pT and pN) should always be reported, if available. Clinical stage should be reported only when pathological stage data are not available.

Data on stage are requested only for the following malignancies: stomach, colon-rectum, skin melanoma, breast, cervix uteri, corpus uteri, ovary, prostate, testis, kidney, bladder and thyroid.

3.15.1 TNM[5] (seven-character variable) **Optional**

Pathological codes for tumour size (pT) and nodal involvement (pN) should be used if available. Clinical or pathological codes for metastases (M) are both acceptable.

The T, N and M components each comprise a single digit, which may be followed by one or two letters (T) or just one letter (N, M). As an example, an *in situ* breast cancer would be coded:

T = _ i s
N = 0 _
M = 0 _

where "_" indicates a blank space. A more advanced breast cancer might be coded:

T = 1 a _
N = 1 b
M = 0 _

3.15.2 Condensed TNM (three-digit variable) **Optional**

When T, and/or N, and/or M codes have not been explicitly recorded in the clinical/pathological records, the cancer registry should attempt to score extent of disease according to the Condensed TNM:

Condensed T

1 = **Localised.** This category comprises T1-2 tumours.
Exceptions: T3 tumours of the thyroid, breast, melanoma (see below), and eye (except sarcoma of orbit) which are also to be considered localised.
2 = **Advanced.** This category comprises T3-4 tumours.
Exceptions: T2 tumours of ovary, fallopian tube, placenta, bone and soft tissues, which are to be considered advanced.

Special condensed T codes for melanoma of the skin
Code as **localised** (condensed T1) those melanomas in TNM categories T1-3 (corresponding to Breslow thickness less than or equal to 4.00mm and Clark levels II-IV)
Code as **advanced** (condensed T2) those melanomas in TNM category T4 (corresponding to Breslow thickness greater than 4.00mm and Clark level V).

Condensed N

0 = No regional lymph-node metastases
1 = Metastasis in the regional lymph-nodes

Condensed M

0 = No distant metastasis
1 = Distant metastasis

3.16 Size of tumour in millimetres (three-digit variable) **Optional**

This should represent the maximum tumour diameter, and should be based on histological examination, if available.

3.17 Number of examined nodes (three-digit variable) **Optional**

Report the exact number of lymph nodes examined, as recorded in the pathological records. If no information is available, or if pathological examination was not performed, code 999.

3.18 Number of metastatic nodes (three-digit variable) **Optional**

Report the exact number of lymph nodes reported as containing tumour, as recorded in the pathological records. If no information is available, or if pathological examination was not performed, code 999.

3.19 'C' (certainty) factor (one-digit variable) **Optional**

The certainty factor reflects the likely validity of the disease stage data for a given case, in relation to the diagnostic methods used to determine it. It refers to the diagnostic examinations carried out to detect or exclude local extension and distant metastases. A simplified classification should be used, as proposed by the ENCR Working Group on extent of disease (http://www.encr.fr/).

1 = C1 - evidence from standard diagnostic methods only
2 = C2 - evidence from special diagnostic methods

Use code 2 (C2) when computerised [axial] tomography (CT or CAT) scan, ultrasonography, nuclear magnetic resonance (NMR) imaging (MRI), or surgery have been used to explore the anatomic region where the tumour is located for cancers at the following sites (ICD-O-3 codes):

- head and neck (C00.X- C14.X and C30.X-C32.X)
- thorax and mediastinum (C15.X, C33.X-C38.X): examinations may also include mediastinoscopy
- pelvis (C51.X-C63.X, C67.X): examinations may also include laparoscopy

Exceptions:

- digestive tract (C16.X-C25.X): use of code C2 requires evidence of liver imaging
- breast (C50.X) and prostate (C61.X): use of code C2 requires evidence of bone imaging (scintigraphy or multiple X-ray investigation)

where "X" indicates any of the valid fourth digits in ICD-O-3, or a blank.

Note: The C code should be used to indicate whether the relevant examination has been performed, regardless of the result of the examination (positive or negative).

3.20 Treatment **Optional**

First course of therapy after diagnosis

3.20.1 Surgery with curative intent (one-digit variable)

1 = Yes
2 = No
9 = No information

3.20.2 Chemotherapy with curative intent, including adjuvant (one-digit variable) **Optional**

1 = Yes
2 = No
9 = No information

3.20.3 Radiotherapy with curative intent, including adjuvant (one-digit variable) **Optional**

1 = Yes
2 = No
9 = No information

3.20.4 Other therapy with curative intent (one-digit variable) **Optional**

For example, hormonal treatments (tamoxifen, etc), and targeted treatments (monoclonal antibodies):

1 = Yes
2 = No
9 = No information

3.20.5 Symptomatic treatment (one-digit variable) **Optional**

For example, radiotherapy given to bone metastases, or intestinal deviation:

1 = Yes
2 = No
9 = No information

3.21 Underlying cause of death (3- or 4-character code) **Optional**

Some registries have access to the underlying cause of death. This enables comparison of cause-specific and relative survival. Registries should specify in an accompanying letter whether the cause of death was coded by cancer registry personnel or by the regional or national vital statistics office.

ICD-9 or ICD-10 are allowed.

4. Additional data

4.1 Life tables

Life tables representing the background mortality in the general population of the territory covered by the cancer registry must be provided, as for EUROCARE-4. Registries that participated in previous cycles of EUROCARE are requested to send an update from 2003 to the most recent year available.

Registries participating in EUROCARE for the first time should send life tables covering their entire period of diagnosis and follow-up. The calendar periods must be indicated in the file.

All-cause mortality rates in the general population, by sex, age and calendar year, should be provided to 6 decimal places or an equivalent number of significant figures (e.g. 0.012345 for a rate of 1,234.5 per 100,000). Since general population mortality is highly dependent on age, mortality rates should preferably be given in one-year age classes: If this is not possible, age classes should not exceed five years: in this case, please specify how the life tables were smoothed. It is essential to have accurate mortality data for the elderly in order to be able to estimate relative survival accurately in this age group. The oldest age class can be open (e.g. 90 years and over), but the lower boundary of the oldest age class should not be less than 85 years.

The source of demographic data should be documented. Cancer registries should also provide the name of a **reference person** who can provide more detailed information on demographic and mortality data in the general population, in the event of queries.

4.2 Population data

Data on the population of the area covered by each registry are required for computing prevalence and for checking data completeness. Registries should supply tables of the number of inhabitants in the registry area by age (5-year age classes), by sex and for each calendar year.

5. File format and submission

A secure, web-based procedure for uploading survival data to the EUROCARE data analysis centre (Rome) is being set up. Instructions will be circulated to the Working Group soon.

Life tables, population data and an accompanying letter should also be sent by surface mail to:

Dr Riccardo Capocaccia
Istituto Superiore di Sanità
Surveillance and Health Promotion
Cancer Epidemiology Unit
Viale Regina Elena 299
I-00161 Rome
ITALY

riccardo.capocaccia@iss.it

EUROCARE-5 DATA FORMAT

The structure of the record format for EUROCARE-5 is shown in the table below. All 26 fields listed in the table must be included in the record, separated by the character @. Variables for which data are not available for a given tumour, or optional variables that are not being supplied, should be represented by null fields (blank) between two separators.

Field	Variable	No. of characters	Protocol item
1	IARC check flag	1	3.1
2	Sex	1	3.2
3	Date of birth	6	3.3
4	Date of diagnosis	8	3.4
5	Date of tumour registration	8	3.5
6	Date of last known vital status	8	3.6
7	Vital status	1	3.7
8	Primary tumour site	3	3.8
9	Microscopic confirmation of diagnosis	1	3.9
10	Morphology and behaviour	5	3.10
11	Summary extent of disease at diagnosis	1	3.11
12	Patient identification code	10	3.12
13	Tumour identification code	10	3.13
14	Multiple tumour code	2	3.14
15	TNM	7	3.15.1
16	Condensed TNM	3	3.15.2
17	Size of tumour in millimetres	3	3.16
18	Number of examined nodes	3	3.17
19	Number of metastatic nodes	3	3.18
20	‘C’ factor	1	3.19
21	Surgery with curative intent	1	3.20.1
22	Chemotherapy with curative intent	1	3.20.2
23	Radiotherapy with curative intent	1	3.20.3
24	Other therapy	1	3.20.4
25	Symptomatic treatment	1	3.20.5
26	Underlying cause of death	4	3.21

References

1 Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J, eds. *Survival of cancer patients in Europe: the EUROCARE study (IARC Scientific Publications No. 132)*. Lyon: International Agency for Research on Cancer; 1995

2 Berrino F, Capocaccia R, Estève J, Gatta G, Hakulinen T, Micheli M, Sant M, Verdecchia A, eds. *Survival of cancer patients in Europe: the EUROCARE-2 study (IARC Scientific Publications No. 151)*. Lyon: International Agency for Research on Cancer; 1999

3 Berrino F, Capocaccia R, Coleman MP, Estève J, Gatta G, Hakulinen T, Micheli A, Sant M, Verdecchia A, eds. *EUROCARE-3: the survival of cancer patients diagnosed in Europe during 1990-94*. *Ann Oncol* 2003; 14 (Suppl. 5): pp1-155

4 Capocaccia R, Gavin A, Hakulinen T, Lutz JM, Sant M (eds.). *Survival of cancer patients in Europe, 1995-2002. The EUROCARE-4 Study*. *Eur J Cancer* 2009; 45

5 Leslie H, Sobin LH et al. *TNM classification of malignant tumours. 6th ed*. New York: John Wiley & Sons, 2002