Protocol for updating the cancer registry-based European High Resolution Project

13 December 2013

INDEX

1. INTRODUCTION

European High resolution (HR) studies on breast, colorectal, prostate, testis and stomach cancer were carried out in the framework of the past EUROCARE-2 and 3, demonstrating that at least some European Cancer Registries (CRs) are able to collect clinical information on stage, staging procedures, diagnostic exams, treatment and follow-up, additionally to the routinely collected CR data. Since these first HR experiences – documented in many scientific publications [1-15] - HR studies were initiated in many countries, aimed to explain the reasons of the survival differences between regions of a same country, and to investigate patterns of cancer care and adhesion to guidelines.

On November, 6th 2012 a European high resolution study workshop was organised in Italy, at Milan Malpensa airport, in the framework of the European Partnership for Action Against Cancer (EPAAC): cancer registries from 11 EU countries participated, together with representatives of the European Network of Cancer Registries (ENCR) and of the Joint Research Centre (JRC, European Commission). The participants expressed a shared interest to work for the development of an EU HR project as a group, highlighting the need to involve clinicians in the study process. Following this first workshop, the HR study activity proceeded with further meetings and contacts between researchers and clinicians, to agree on a final study protocol and to start the study. The present protocol version includes the comments of participants to the second HR workshop of September, 25th-26th 2013, hosted at JRC, Ispra, and of November 21st 2013, hosted at Fondazione IRCCS Istituto Nazionale dei Tumori, Milan.

2. AIMS

As in the past, the final aim of HR studies is **to help interpreting differences in cancer outcome and survival** across European populations, using more detailed data on diagnosis, treatment and follow-up than those usually available in the routine activity of cancer registries.

However, for most cancers, this objective can be achieved only after a considerable time lag between data collection, survival analyses and publication of results, while to timely capture changes and improvements in diagnosis and treatment, it is necessary to collect data of patients diagnosed as recently as possible. In this view, collection of relevant clinical data, in addition to those routinely registered would be more effective and sustainable than the traditional high resolution studies based on retrospective sampling and data retrieval. However, this means to wait for an adequate length of follow-up, before survival analyses can be carried out. Thus, short and medium or long term HR study aims can be envisaged.

SHORT TERM AIMS of the new HR studies:

- To investigate and compare patterns of cancer care and adherence to internationally agreed clinical guidelines for diagnosis and treatment across countries, regions, or groups of patients. In the next 2-3 years European projects (e.g. "European Guide on quality improvement in Comprehensive Cancer Control", CANCON) will be implemented, aimed to harmonise existing national clinical guidelines, and develop European guidelines for at least some cancers. A network of population-based cancer registries carrying out HR studies, will greatly help monitoring adherence to these guidelines across countries and assess their impact on current clinical practice;
- To investigate specific short term cancer outcomes, e.g relapse and disease free interval; short-term (≤1-year survival can be meaningful for some cancers, e.g. lung cancer); diffusion of new targeted treatments in the current clinical practice and policies of using them; outcomes of interest can be identified over the study time-frame;
- To increase the quality of data collected by CRs on stage and morphology and to better characterize the cancers, by means of validated biomarkers (e.g. HER2 over expression and KI67 for breast cancer); updated morphology classification for haematological malignancies). A collaboration will be pursued with the molecular pathobiology Working Group of the European Society of Pathologists and with the Organisation of European Cancer Institutes (OECI), to study prognostic and predictive molecular biomarkers which could be associated with outcome and could be used in epidemiological studies;
- To test the feasibility of collecting data on co-morbidity potentially influencing cancer prognosis. The ability to manage a patient with cancer successfully very much depends on the general health of the patient, thus variations in care and outcome may depend on the presence of co-morbidity and on his/her functional status;
- To set the basis for updating life status and clinical follow-up of cases included in the HR studies, for future investigations.

MEDIUM AND LONG TERM AIMS of the new HR studies:

- To update follow-up for life status and relapse, after an adequate length of time since diagnosis, e.g. 5 or more years. For this purpose, in addition to the usual procedures adopted by CRs to update life status and relapse, the linkage of the HR records with those included in the EUROCARE basic database will be carried out. The linking procedure is facilitated by the variables common to both HR and basic databases. For this reason, the patient identification code

used in the HR studies must be the same of that in the EUROCARE basic database, or the algorithm for decrypting id codes should be made available;

- To analyse overall and disease-free survival, in relation to patterns of care and other information previously collected. The results of these analyses will contribute to explaining the reasons for survival differences between regions and group of patients.

3. STUDY DESIGN

Most participants in the HR study workshops expressed support towards a **prospective data collection**, which will include cases diagnosed more recently, typically those cases diagnosed in the latest or current year of registration. As cases are registered to be included in incidence series, additional information on care items can be added at the same time to the variables usually recorded by the registry. This approach is useful to obtain timely information on patterns of care and allows for descriptive analyses on the frequency of procedures or adhesion to guidelines. In order to have an adequate follow-up for studying survival, the evaluation of survival would be postponed compared to past HR studies, which recruited cases retrospectively. However, also the 1- or 2-year follow-up potentially available with a prospective approach can be informative for some cancer sites (e.g. lung cancer) and for certain outcomes (e.g. relapse).

The survey carried out amongst the CRs participating in the HR workshops indicated that most of them are now collecting cancer cases incident after the year 2010, thus, 2011-2013 could be a suitable study period for additional HR data collection, common to all registries and sufficiently recent to detect recent changes in cancer care. However, CRs which would like to participate in this project, but are collecting cases incident before 2011 or incident in 2014, are also invited to submit data: their data will be evaluated for specific analyses.

The relevant files will be archived for future investigations and for updating follow-up (life status, relapse, and subsequent tumours) of these cases, in due time.

As for the previous HR studies, all the relevant sources of information (clinical notes, ambulatory and hospital files, mortality files etc.) are to be considered to reconstruct the clinical history. Each registry will identify the most appropriate source of information to describe patterns of care and follow-up in their area.

3.1 Cancers to be included

Based on criteria of high public health and scientific significance, as well as current and past HR experiences, during the HR study workshop of September, 25th -26th 2013, it was agreed to include the following cancer sites in the new HR study: breast, colorectal, lung, skin melanoma, and lymphoid malignancies. These cancer sites are those most commonly included in the HR studies currently ongoing in European cancer registries and on which a large participation can be envisaged.

3.2 Inclusion criteria

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 [16]. Both microscopically verified and non-verified cases must be included. In particular, the inclusion criteria for each cancer under study are:

Breast: adult (≥15 years) women cases diagnosed with in situ (/2) or malignant (/3) breast cancer (ICD-O-3 topography code: C50)

Colorectum: adult (≥15 years) cases diagnosed with malignant (/3) colon (C18), rectum (C20), rectosigmoid junction (C19) and anus and anal canal (C21) cancers

Lung: adult (≥15 years) cases diagnosed with malignant (/3) lung (C34) cancer

Skin melanoma: adult (≥15 years) cases diagnosed with malignant (/3) skin melanoma (C44)

Lymphoid malignancies: adult (≥15 years) cases diagnosed with follicular (ICD-O-3 morphology codes: 9690-9691, 9695, 9698) and diffuse large B-cell (9678-9680, 9684) lymphoma.

3.3 Sampling procedure

Taking into account that the size of the registration area is different among European CRs and in order to draw unbiased samples with respect to the whole incidence series, using a "prospective" data collection, the study population for each CR participating to the study may include one of the following sampling procedures:

- one (or more) year of complete incidence, for CRs covering a small-medium area
- an administratively defined sub-area (e.g. one region in a country, one health district in a region that might include urban, peri-urban and rural areas), for CRs covering a large area (e.g. national CRs)
- six months of incidence for CRs covering a large area (e.g. national CRs)

All incident cases occurring in the relevant period/area are to be included. The completeness of cases included in the study can be checked only after completion of the whole incidence data regardless the method of sampling.

3.4 Study dimension

Past HR studies required from 100-200 [1-4,7,8,12-14] to 500 [5,9,10,15] cases per cancer site per registry, depending on study period and incidence level in the registry's area. In past HR studies on breast cancer, 5-year survival differences resulted statistically significant with hazard ratio figures of 1.60 (among European countries) or up to 1.40 (among Italian areas).

For the present HR study, the number of cases necessary for detecting significant differences was calculated by proper procedures, depending on the outcome (e.g. odds of being treated according to a given procedure and 5-year survival) by taking into consideration the past experiences.

In particular, assuming survival estimates for European adult cancer patients diagnosed in 2000-07 [17] as reference, 500 cases are required to detect at 5% level of significance and 80% power significant differences in the survival range 74-85% (or, equally, to detect 30% variations – reduction or increasing – in the risk of death) for breast cancer; for colorectal cancer in the range 52-66% (or, equally, 20% variations in the risk of death); for lung cancer in the range 7-13% (or, equally, 15% variations in the risk of death); whereas 300 cases are required to detect significant differences in the survival range 72-87% (or, equally, 35% variations in the risk of death) for skin melanoma, and 200 cases in the range 48-70% (or, equally, 30% variations in the risk of death) for lymphomas.

Moreover, the estimated sample sizes (i.e. 500, 300, and 200, respectively) are sufficient to detect odds ratios (ORs) of 1.5, 2 or 2.3 (or, equally, odds ratios of 0.55, 0.5 or 0.4, respectively) as significant (at the 5% level of significance level and 80% power), after assuming that: (i) the probability of receiving the proper standard care in the reference group is at least 50%, (ii) more than one variable will be included in the logistic regression model, and (iii) moderate ($\leq$0.5) correlations between variables included in the models exist. We note that the estimated survival ranges and ORs are in line, respectively, with survival ranges found in Europe in 2000-07 for each analysed cancer site and with ORs found in the past breast and colorectal cancer studies [5,9].

Thus, in the present study, based on the past HR experiences, a minimum of 500 cases per registry is requested for breast, colorectal, and lung cancers, while a minimum of 300 cases is requested for skin melanoma and a minimum of 200 cases for lymphoid malignancies, among those incident in the chosen year and/or area.

If with the chosen method of sampling, 500 (300 or 200, depending on the cancer site) cases are not recorded during the selected incidence period, the CR should extend its sampling (maintaining the same sampling procedure) forward/backward in time, in order to reach the required number of cases. However, for diffuse and follicular lymphomas, taking into account cases incident in some small areas covered by cancer registration (<50 cases/year), at least two complete years of incidence are requested.

3.5 Cancer registries involved

All European CRs are invited to participate in this study. The procedures developed in past HR studies will be carried out to check the quality and completeness of data to be included in the final analyses.

3.6 Data collection

3.6.1 List of variables

The list of variables (see Annex A and Annex B) to be collected reflects the decisions taken during the three HR workshops held in Malpensa on November, 6th 2012, at the JRC-Ispra on September 25th-26th 2013, and at the Fondazione IRCCS Istituto Nazionale dei Tumori on November 21st 2013, and contains variables collected in past HR studies in several countries. This list is composed by a set of variables common to all cancers, and a set of additional cancer specific variables. The present study proposal includes cancer specific variables for breast, colorectal, lung, skin melanoma and some lymphoid malignancies, i.e. the first cancers for which an agreement has been reached during the HR workshops.

In order to facilitate the record linkage and integration of the HR studies with the EUROCARE general survival study, a number of variables are common between the two studies. In order to harmonise data used for different cancer registry projects, the definition and coding of core variables such as date of diagnosis, stage, base of diagnosis, multifocality follows the ENCR-EUROCARE rules and the indications of the ENCR-EUROCARE workshop of October, 15th 2013 at JRC, Ispra.

3.6.2 File format

Data for <u>each cancer site</u> could be collected according to two different formats:

1. HR **text files** (.txt) should be formatted as follows (similarly to what is requested for EUROCARE and ENCR data):

- one record per line

- delimited field format (; or | should be used as separators)
- pre-defined format and record layout (as reported in the attached file including the list of variables). Variables for which data are not available or optional variables that are not being supplied should be represented by null fields among the two separators if for those variables was not specified a specific code for missing information.

2. HR **Access databases**, prepared and distributed to the participating CRs by the Analytical Epidemiology and Health Impact Unit, Fondazione IRCCS Istituto Nazionale dei Tumori.

A single file for each cancer site should be uploaded by each CR. Specific instructions for each uploading modality, will be circulated.

4. DATA SUBMISSION, RESOURCES AND CHECKS

Each European cancer registry will be able to transfer its data securely and efficiently by using a dedicated FTP server hosted by the Fondazione IRCCS Istituto Nazionale dei Tumori in Milan (INT), accessible by a username and password centrally assigned by the Analytical Epidemiology and Health Impact Unit (coordinating centre) at INT. Instructions will be circulated to the participating European CRs in January 2014.

For the phases pertaining to uploading data, quality checks, data correction, pre- and basic analyses, the HR files will be centralised at the coordinating centre at INT.

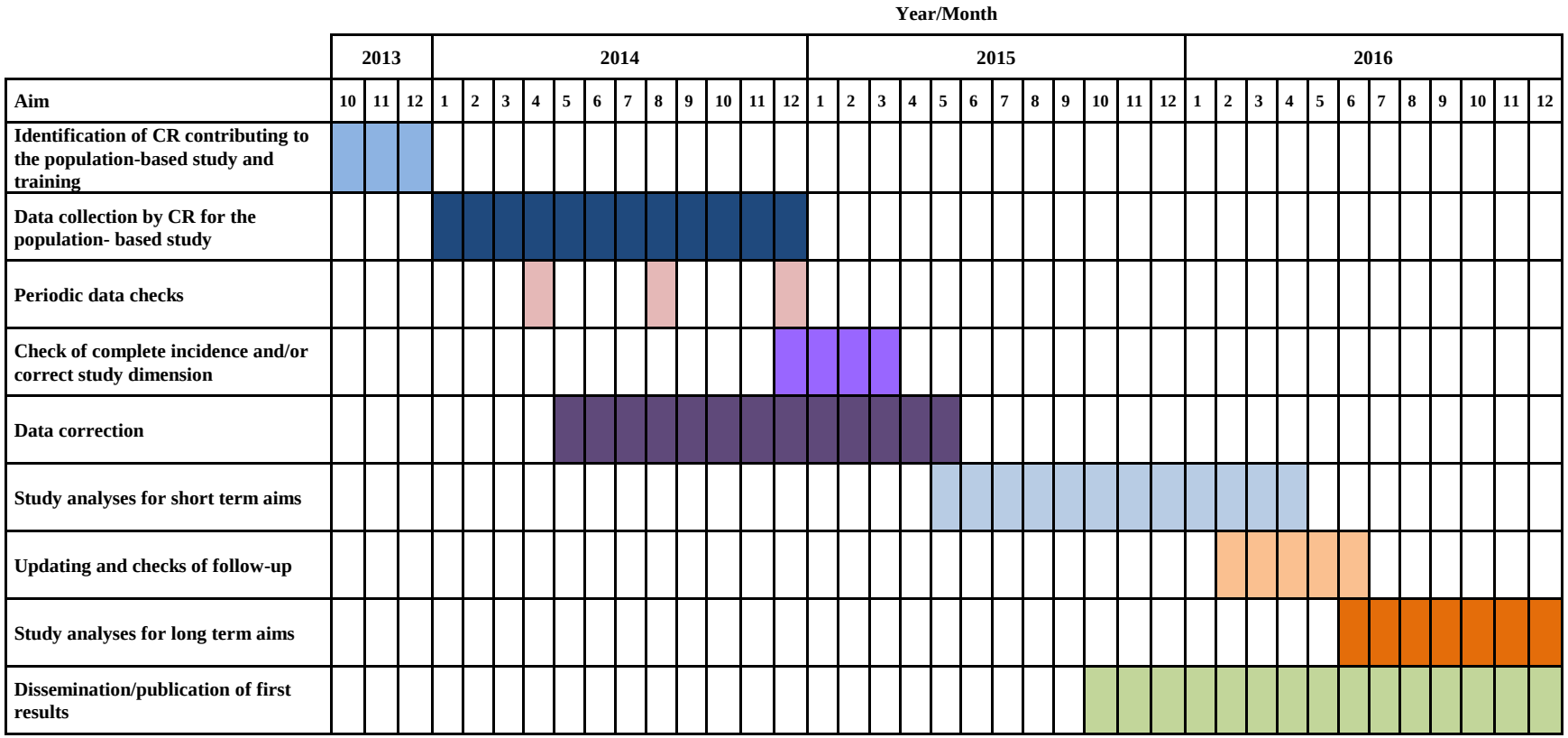
Personnel of this unit will merge, clean and store CR data in a unique database, with methods similar to those used for the basic EUROCARE study. Systematic checks of data will be performed periodically during the phase of data collection (every 4-6 months), to correct errors and solving problems timely. Records with errors or incongruence will be sent back to the CRs for correction.

The completeness of data with respect to incidence and survival series will be checked after the conclusion of HR data collection. For this aim, as well as for reaching medium and long term aims (life tables, population data and source of demographic data for estimating relative or net survival) resources are available at the Cancer Epidemiology Unit, Istituto Superiore di Sanità (ISS), Rome.

A linkage between the HR databases and the EUROCARE data to update vital status and date at the last known contact is envisaged for future survival analyses.

5. TENTATIVE TIMETABLE

In the short term, the main output of the European High Resolution study will be the analysis of patterns of care and adhesion to internationally agreed clinical guidelines. In the medium-long term, the main output will be the survival (relative, net or disease-free) analysis in relation to patterns of care. By January 2014, the participating cancer registries will start collecting data on patients diagnosed from January, 1st 2011 to December 31st 2013. During 2014 and 2015, these data will be checked and analysed.



6. AUTHORSHIP, PUBLICATION POLICY AND DATA AVAILABILITY

For this new European HR study, the EUROCARE publication rules will be basically adopted, unless proposals for changes are received.

The authorship of scientific articles, in order to acknowledge each collaborator contribution, should include:

First author: study design, responsibility of analyses, writing the paper

One or two authors of the Analytical Epidemiology and Health Impact Unit of the INT: data managing, quality checks, basic or complete statistical analyses.

Other contributors to the study, e.g. clinicians, external collaborations, etc.

One author for each registry contributing data. If the journal to which the paper is submitted only allows a restricted number of authors, a HR working group will be named, with all contributors not included amongst the named authors.

As regards to the publication plan, two different phases are scheduled.

Phase 1: the first basic HR analyses will be centralised at the coordinating centre (Analytical Epidemiology and Health Impact Unit, at INT) and related reports with summary results will be ready by late 2015 or during 2016.

Phase 2: Further analyses and publications will be discussed and scheduled with all CRs involved in the study. During this phase, the possibility and the procedure of data releasing to researchers willing to carry out specific studies will be discussed and formalised.

7. CONFIDENTIALITY AND STORAGE

Cancer data will be stored individually, but anonymously (anagraphical information will not be recorded, but alphanumeric codes assigned by each cancer registry allow differentiating cancer patients by each other), in a dedicated server hosted by the Fondazione IRCCS Istituto Nazionale dei Tumori not connected to the web.

8. FUNDING PERSPECTIVES

The HR data collection requires a considerable amount of resources and extra work load. Some registries are committed and funded by the government or other national bodies for collecting data on patterns of care, and the data collected in this frame could be included in a European HR study. However, at the moment, no resources are immediately available for funding a European collaborative HR project, thus presently the collection of data for the HR studies should be supported by local initiatives and scientific societies able to provide funds. Resources for analysis

and coordination (personnel, computers, limited funds for meetings and travel) are available at INT and ISS, through national and EU projects.

Within the EPAAC project a collaboration has been established between WP8 (optimizing EU research funding) and WP9 (cancer information and data). The high resolution studies will represent a specific task of a draft pilot project on "outcome research", to be further developed and submitted to the EU, in the perspective of future funding.

9. CONTACTS

For any questions, please contact

Dr. Milena Sant (coordinator)

Tel.:+39-02-23903519-2902-2867; E-mail: milena.sant@istitutotumori.mi.it

Dr. Pamela Minicozzi

Tel.:+39-02-23903520; E-mail: pamela.minicozzi@istitutotumori.mi.it

EUROCARE Secretariat (Chiara Margutti, Camilla Amati)

Tel.:+39-02-2390-2902-2869; E-mail: eurocare.secretariat@istitutotumori.mi.it

Analytical Epidemiology and Health Impact Unit

Department of Preventive and Predictive Medicine

Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

10. REFERENCES

1 Sant M, Eurocare Working Group. Differences in stage and therapy for breast cancer across Europe. Int J Cancer 2001;93:894-901.

2 Sant M, Allemani C, Capocaccia R, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416-22.

3 Sant M, Allemani C, Berrino F, et al. Breast Carcinoma Survival in Europe and the United States. Cancer 2004; 100:715-22.

4 Allemani C, Sant M, Berrino F, et al. Prognostic Value of Morphology and Hormone Receptor Status In Breast Cancer – A Population-Based Study. Br J Cancer 2004;91:1263-8.

5 Allemani C, Storm H, Voogd A, et al. Variation in "standard care" for breast cancer across Europe: a EUROCARE-3 high resolution study . Eur J Cancer 2010;46:1528-36.

6 Allemani C, Sant M, Weir HK, et al. Breast cancer survival in the US and Europe: a CONCORD high-resolution study. Int J Cancer 2013;132:1170-81.

7 Gatta G, Capocaccia R, Sant M, et al. Understanding variation in survival for colorectal cancer in Europe: a EUROCARE high resolution study. Gut 2000;47:533-8.

8 Ciccolallo L, Capocaccia R, Coleman MP, et al. Survival differences between European and US patients with colorectal cancer: role of stage at diagnosis and surgery. Gut 2005;54:268-73.

9 Gatta G, Zigon G, Aareleid T, et al. Patterns of care for European colorectal cancer patients diagnosed 1996 – 1998: A EUROCARE high resolution study. Acta Oncol 2010;49:776-83.

10 Bouvier AM, Minicozzi P, Grosclaude P, et al. Patterns of adjuvant chemotherapy for stage II and III colon cancer in France and Italy. Dig Liver Dis 2013;45:687-91.

11 Allemani C, Rachet B, Weir HK, et al. Colorectal cancer survival in the USA and Europe: a CONCORD high-resolution study. BMJ Open 2013;3:e003055.

12 Gatta G, Zigon G, Buemi A, et al. Prostate cancer treatment in Europe at the end of 1990s. Acta Oncol 2009;48:867-73.

13 Sant M, Aareleid T, Artioli ME, et al. Ten-year survival and risk of relapse for testicular cancer: A EUROCARE high resolution study. Eur J Cancer 2007;43:585-92.

14 Lepage C, Sant M, Verdecchia A, et al. Operative mortality after gastric cancer resection and long-term survival differences across Europe. Br J Surg 2010;97:235-9.

15 Bouvier AM, Sant M, Verdecchia A, et al. What reasons lie behind long-term survival differences for gastric cancer within Europe? Eur J Cancer 2010; 46: 1086-92.

16 Fritz A, Percy C, Jack A, et al. International classification of disease for oncology (ICD-O), 3rd edn. World Health Organization, Geneva, 2000.

17 De Angelis R, Sant M, Coleman MP, et al. Cancer Survival in Europe 1999-2007 by country and age: results of EUROCARE-5 – a population-based study. Lancet Oncol 2013; in press.