IPAAC – TASK 7.3: PILOTING THE INTEGRATION OF DATA ON CANCER COSTS
Task Leader: Silvia Francisci ISS

PROTOCOL FOR COLLECTING POPULATION-BASED CANCER PREVALENCE DATA INTEGRATED WITH OTHER AVAILABLE ADMINISTRATIVE SOURCES OF HEALTH CARE AND COSTS DATA IN EUROPE

DECEMBER 2018

INTRODUCTION

Describing and identifying costs related to specific health care needs provide useful information to better allocate health care resources and facilities, and to reduce possible inequalities among cancer patients within and between countries.

In Italy, the EPICOST study has recently been concluded. It aimed at estimating the number of cancer patients in Italy and corresponding healthcare expenditures according to a three-phase disease pathway, from diagnosis to possible recovery or death, and to evaluate cost-effectiveness of selected health care interventions.

The study used individual level information collected from population based cancer registries (CR) and integrated with the following administrative data sources:

   o   Hospitalizations: Hospital Admissions database

   o   Ambulatory/outpatient visits: Outpatient Services database

   o   Drug prescriptions: in hospital, in ambulatory, in pharmacy

AIM OF THE PILOT STUDY - TASK 7.3
This pilot study will assess the feasibility of **extending to other European countries the procedures developed in the framework of the EPICOST Italian project,** integrating data from population-based Cancer Registries (CRs) and other administrative data sources, in order to estimate costs directly related to diagnosis, care and follow-up of cancer patients (cost profiles).

The task leader and partners will derive procedures to reconstruct cancer–specific cost profiles that will be shared among MSs participating to this pilot.

INCLUSION CRITERIA

Prevalence cohort
Invasive, primary, malignant neoplasms of **rectum (ICDO3 topography C19- C20), colon (ICDO3 topography C18), pancreas (ICDO3 topography C25), and skin melanoma (ICDO3 topography C44, morphology 8720-8790) in adult population (aged 15+)** are eligible for inclusion in the pilot study- task 7.3 (these are called **index tumours**).

The study cohort includes patients diagnosed with the index tumours during all years of activity of the CR and still alive at prevalence date (**prevalence cohort**). The prevalence date is the last

updated available according to the CR database and an entire year of life status follow-up after the prevalence date must be available for the prevalence cohort. In case of prevalent cases with multiple primaries, only the following cases will be included: a) patients with index tumours diagnosed as most recent; **and** b) if the other primaries (any cancer type) occurred 5 or more years before the index tumour diagnosis date.

The prevalence study cohort is to be established by the single Cancer Registry or it could be centrally selected by using the EUROCARE-6 database (if it included the last available and updated CR data).

Administrative/health care data sources

At least cancer and not-cancer related hospitalisation data

DIAGNOSTIC CLASSIFICATION
Anatomic site, tumour morphology and behaviour must be coded according to the International Classification of Diseases for Oncology (ICD-O-3), published in 2000 and updated in 2011.

DATA SOURCES AND DATA LINKAGE
The pilot study-task 7.3 uses data **at individual level linked by the CR to different administrative/health care data sources and to the mortality file,** in order to reconstruct patterns of care and corresponding costs of study cohort patients **in a 3-year period** spanning from 2 years before the prevalence date to one year after the prevalence date (**study period**).

All records identified from the record linkage as related to patients of the prevalence cohort **within the study period** are provided by the CR. The CR must provide **a code allowing the tracing of the single patient of the prevalent study cohort in all data sources provided (patient-ID)**.

Number and contents of data sources considered for the linkage might vary according to the country health care data system. The aim is to include as much information as possible to estimate the total amount of health care expenditures directly related to diagnosis, treatment and monitoring of the prevalence cohort during the study period.

A questionnaire investigating the availability and contents of data sources provided by the CR participating to the tasks included in the IPAAC-WP7 will be administered by the task 7.1 by September 2018 (according to the minutes of the WP7 Kick-off meeting held in Milan May 31st 2018).

INFORMATION REQUIRED

From Cancer Registry database: all variables included in the 2015 ENCR-JRC Call for Data study protocol. The dataset will include a record for each patient included in the prevalence cohort and for each tumour (i.e. patients with multiple tumours will have multiple records).
Information on stage at diagnosis is requested only for prevalent cases diagnosis up to 12 months before the prevalence date.
For each individuals all tumours occurred before the prevalence date should be provided.

From administrative/health care data sources:
- ➢ Patient-ID (the same one used in the CR database sent for the pilot)
- ➢ Information on:
    - ➢ Type of procedure (diagnostic procedures, treatments, outpatient procedures and visits) classified according to the ICD9-CM (ICD10-CM ??), pharmaceutical prescriptions classified according to the Anatomical Therapeutic Chemical Classification System (ATC code), other (please specify if you have suggestions)
    - ➢ Date of procedure
    - ➢ Quantity of procedure
    - ➢ Cost per unit of procedure (in Euros)
    - ➢ Total cost of procedure (in Euros)
    - ➢ Additional variables are data source- specific: for example, in the case of the Hospital Discharge database: regimen (with or without overnight stay in hospital), number of days of stay, multiple diagnostic codes (main diagnosis, secondary diagnoses up to …), multiple treatment codes (main treatment, secondary treatments up to …), DRG code, …

The datasets (one for each health care source) will include a record per patient included in the prevalence cohort and per procedure (i.e multiple procedures for the same patient correspond to multiple records).

From the mortality file:

- ➢ Patient-ID (the same one used in the CR database sent for the pilot);
- ➢ Date of death;
- ➢ Cause of death

LIST OF INDICATORS

Cost indicators

Each patient of the prevalence cohort contributes to the study with a 12-month time interval. We defined three phases of care: initial, continuing and final. On prevalence date, each patient of the prevalence cohort belongs to only one phase of care, depending on the interval between prevalence date and diagnosis date and on the occurrence of death for cancer during the following year.

Each case is linked to the available administrative/health care databases in order to trace every event of interest during the follow up time. In order to take into consideration only those events that are related to the index tumours, a list of events (specific for each index tumour) is utilised for each database. These lists were created by expert oncologists in the framework of the EPICOST

Italian project and referred to diagnoses, interventions and procedures coded according to the ICD9-CM classification for Hospital Discharges DB and for outpatient services DB, and to the ATC classification system for drug prescriptions DB. The lists are available for colon and rectum cancers only.

Costs are expressed in Euros and are defined as the direct expenditure related to diagnosis, treatment, follow-up and end-of-life care provided to cancer patients.

The following indicators are considered for each of the three phases of care:

Patient monthly average cost Ci: all costs sustained on average for a patient in month i, obtained by dividing costs sustained for all patients in month i by the corresponding number of person-months.

Patient annual average cost CA: all costs sustained on average for a patient in a year, obtained by summing up patient monthly average costs, i.e. $C_A = \sum_{i=1}^{12} C_i$

A **cost profile** is a series of 36 patient monthly average costs Ci over the three phases of care.

Total annual cost: all costs sustained in 12 months for all patients, obtained by multiplying the patient annual average cost CA by the total number of patients. These costs are computed by phase of care and/or by type of health care service.

We identify homogeneous groups of patients according to clinical and demographic variables affecting the patterns of care: age, stage at diagnosis (for the initial phase only) and compute costs as simple averages over patients belonging to the same homogeneous group.

Patterns of care indicators

In order to better describe and interpret results on costs in the initial phase of care, a list of patterns of care indicators is computed, the list is specific for each index tumour considered. Here an example of indicators computed by age at prevalence and stage at diagnosis, applicable to colon cancer:

- percentage of patients receiving at least one surgery treatment;
- percentage of patients receiving at least one chemotherapy over all patients in initial phase of care;
- time occurring between surgery and chemotherapy.

DATASETS FORMAT
The datasets (one for each data source) are required in CVS format with semicolon (;) separating the variables.

TEAM OF THE PILOT STUDY - TASK 7.3
The pilot study will involve a multidisciplinary team that, besides CRs, includes statisticians, epidemiologists and health economists. An IT company (TBD) will be subcontracted by ISS

to   provide technical assistance to implement the platform for uploading and checking cancer registries data.

The <u>EXAMPLE</u> of data from administrative/health care data sources used for <u>EPICOST in ITALY</u>

1. Hospital Discharge database (HD),
2. Outpatient Services database (OPS),
3. Drug Prescriptions database (DP)
4. Hospital Drugs (HDP)

1. Hospital Discharge database (HD) record track
   - Patient ID (the same one used in the CR database)
   - Demographic variables (sex, place of residence, date of birth, civil status, education level)
   - Type of admission (ordinary, day hospital)
   - Dates of admission at the hospital and of discharge;
   - Diagnosis (principal diagnosis + the other secondary up to five);
   - Diagnostic and intervention procedures (principal intervention + the other secondary interventions up to five);
   - Dates of diagnostic and intervention procedures;
   - Discharge modality (patient death, ordinary discharge, transfer to other unit same hospital, transfer to other hospital);
   - DRG code
   - Total claim (in Euros)

2. Outpatient Services database (OPS) record track:
   - Patient ID (the same one used in the CR database)
   - Dates of service;
   - Code of Diagnostic and intervention procedure;
   - Description of Diagnostic and intervention procedure;
   - Date of Diagnostic and intervention procedure;
   - Branch of the procedure (numerical code corresponding to homogeneous groups of interventions: diagnostic, visits, radiotherapy, genetic tests, …)
   - Quantity: number of diagnostic or intervention procedure;
   - Tariff: unitary cost per single diagnostic or intervention procedure;
   - Total claim (in Euros): Quantity X Tariff, when quantity is >1

3. Drug Prescriptions database (DP) / Hospital Drugs (HDP) record track:
   - Patient ID (the same one used in the CR database)
   - Dates of pharmaceutical prescription;
   - ATC code;
   - AIC code;
   - Quantity: number of doses indicated in the prescription;
   - Tariff: unitary cost per single dose of drug;
   - Total claim (in Euros) corresponding to the total cost of the prescription (quantity X Tariff)