**Protocol for WP3 - Mortality associated with congenital anomalies**

**WP3 leads:** Judith Rankin (UNEW), Anna Pierini (CNR-IFC)

**Researchers:** Svetlana Glinianaia (UNEW), Michele Santoro (CNR-IFC), Alessio Coi (CNR-IFC)

**1. Background and aim**

There is a large variation in child death rates across Europe; in 2012 the child death rates (age 0-14 years) were 60% higher in the UK and Belgium compared to Sweden, with an additional 10 countries being 30% higher than Sweden.[1](#_ENREF_1) Congenital anomalies are a leading cause of perinatal and infant mortality, especially in developed countries.[2](#_ENREF_2) In 2012, congenital anomalies were associated with about a third of infant deaths in the UK.[3](#_ENREF_3) Further studies are needed to identify potentially preventable causes and to understand the source of the variations in child death rates across Europe, including health inequalities.

Existing studies focus on specific conditions for example Down syndrome,[4](#_ENREF_4),[5](#_ENREF_5) spina bifida[6](#_ENREF_6) and cardiac anomalies,[7-9](#_ENREF_7) but there is a lack of detailed information about the survival of children with other specific congenital anomalies. A recent study did report on a large number of specific congenital anomaly groups,[10](#_ENREF_10) however, it was based on data from one region of England and, as indicated above, the mortality rates are likely to vary considerably across Europe. No analysis comparing the mortality of children with specific congenital anomalies above the age of one across Europe has been published.

The American study by Copeland and Kirby (2007) that compared mortality associated with a congenital anomaly using congenital anomaly registry data versus death certificate data,[11](#_ENREF_11) found that an existing congenital anomaly was not commonly reported as an underlying cause of death in a death certificate. The WP3 study will identify deaths up to 10 years of age via the linkage between data from the EUROCAT congenital anomaly registries and data from mortality databases to determine the contribution of congenital anomalies to mortality up to 10 years of age.

Mortality during the first year of life will be examined using the following categories:

a. early neonatal death (death within the first 0-6 days of life)

b. late neonatal death (death within the first 7-27 days of life)

c. postneonatal death (death from 28 days to under 1 year)

This categorisation is important as, worldwide, 45% of deaths in children under 5 years occur in the neonatal period (Unicef statistics, 2015, <http://data.unicef.org/topic/maternal-health/newborn-care/>) and among neonatal deaths, about 36% and 73% occur in the first day and in the first week of life respectively. Early neonatal deaths occurring during the first 24 hours will also be identified where possible.

Mortality during the subsequent 9 years will also be examined.

In order to facilitate interpretation of infant and childhood mortality amongst live births with a congenital anomaly, it is also necessary to have information on the natural history of all pregnancies with the specified congenital anomaly. For example, if the majority of pregnancies with severe cardiac anomalies result in a termination of pregnancy for fetal anomaly (TOPFA), the survival of neonates with a cardiac anomaly will be much higher than if no pregnancies with a cardiac anomaly result in a TOPFA, on the premise that it is the more lethal cardiac anomalies that result in a TOPFA.

The aim of WP3 is to expand the knowledge on the survival of live born children with congenital anomalies up to 10 years of life and to evaluate the potential effect of prenatal diagnosis on survival and risk factors for survival in Europe, in particular any inequalities in survival.

The **specific tasks** are to:

1. Evaluate the survival of babies with congenital anomalies by selected EUROCAT congenital anomaly subgroups across Europe (UNEW)
2. Investigate whether survival of infants and children is associated with occurrence of a prenatal diagnosis (UNEW)
3. Investigate whether there are geographic variations in survival across Europe for selected EUROCAT congenital anomaly subgroups (CNR-IFC)
4. Investigate the association of risk factors (infant sex, birth weight, gestation length, maternal age, socio-economic status, non-European origin of the parents) and survival (UNEW)

**2. Data linkage, definitions and coding of causes of death - live births only**

**Summary of EUROlinkCAT Work Package on Mortality**

Twenty one EUROCAT registries are responsible for ensuring the linkage of their data on all live births with a major congenital anomaly with locally held mortality data. Following a successful application for ethics approval, where this is required, the linkage with the mortality data will be based on a child’s local unique identifier (e.g. NHS number in England & Wales, national unique ID number in Scandinavian countries and regional unique ID number in Tuscany) and/or by matching a number of child details (names, date of birth etc) which uniquely identify the child. The mortality data sources listed in Appendix 1 will be used for the data linkage by the EUROCAT registries. The data provider will produce frequencies of specified variables in the mortality database, to assess data quality.

Each registry will extract a file of all live born congenital anomaly cases and all other pregnancies resulting in a fetal death from 20 weeks of gestation or TOPFA at any gestational age from their local EUROCAT Data Management Program (EDMP). All live born congenital anomaly cases will then be linked to the data on mortality. The linked data file will contain the following information (where available) coded as detailed below (see Tables 1 and 2). The provider of the mortality data will produce a short report outlining the number of cases that were linked / unlinked and the reasons for non-linkage. The linked data file will be stored securely, either within the local registry or within the organisation doing the linkage (see section 3.1).

The registries will be provided with a set of instructions (syntax script) to create pre-specified tables and perform analysis on the data and the aggregated tables and analytic results will be transmitted to the Central Results Repository (CRR) at Ulster University. No individual case data will be transmitted to the CRR.

***2.1. Years to include***

All births from 1st January 1995 up until 31st December 2014, deaths to 31st December 2015. 12 EUROCAT registries (Belgium: Antwerp, Croatia: Zagreb, Denmark: Odense, Finland, Germany: Saxony-Anhalt, Italy: Emilia Romagna, Italy: Tuscany, Malta, Netherlands: Northern, Spain: Basque Country, UK: Thames Valley, UK: Wessex) have data on births from 1995. The remaining registries have data on births from later dates: 1997 (France: Paris), 1998 (UK: East Midlands, UK: Wales), 1999 (Norway), 2000 (UK: North), 2002 (France: Ile de la Reunion), 2005 (UK: South West, Ukraine: West) and 2007 (Spain: Valencia).

**2.2. *Data for extraction: all cases with a Major Congenital Anomaly***

Cases for extraction will be all cases with a major congenital anomaly recorded in the (*insert name of Congenital Anomaly Registry here*) with births from (*insert date*) to (*insert date*). All cases with only isolated minor anomalies as defined in Chapter 3.2 of the EUROCAT Guide 1.4 and Reference Documents (last update version 20/12/2016 - <http://www.eurocat-network.eu/content/Full%20Guide%201%204%20v5%2020_Dec2016.pdf>) will be excluded. Only cases born to mothers resident in the registry area irrespective of place of birth will be included, thereby excluding births within the registry region to non-resident mothers.

Subgroups of congenital anomalies have to be defined in accordance with the latest EUROCAT classifications which are integrated in the EDMP[12](#_ENREF_12) (see chapter 3.3 of the EUROCAT Guide 1.4).Minor anomalies will not be updated before the end of the EUROlinkCAT project to ensure consistency over time. All registries have to use the most recent version of EDMP (Version 6.10), 11 January 2016.

All cases resulting in a live birth will be identified for linkage with mortality data.

***2.3. Linkage of Cases to Mortality Data***

All live born cases will be searched for in *[insert name of any databases containing information from death certificates]* for any deaths occurring from (*insert date*) to (*insert date*). Searches will be conducted using the child’s local unique identifier (preferably) or by matching a number of the cases’ personal details (names, address, child’s sex, date of birth, birth weight etc.) to uniquely identify the case. In the absence of a child’s unique identifier, we strongly recommend that you do not link cases by a single non-specific variable (e.g. date of birth), these variables should be used along with other unique variables (child’s name, maternal name, residential address etc). All linkages will be characterised by the strength of the linkage, and the data provider (or whoever is doing the linkage) will be asked to produce a ‘’linkage report’’. Some countries have electronic databases of births and then these births are linked to deaths, so a case can be identified. If there is insufficient information to identify the birth then clearly the death would also not be identified. Any information concerning the strength of linking or not linking will be collected.

It is extremely important that for all cases that have not been linked to mortality data, the registries or mortality data providers confirm whether the case is alive or whether no linkage could occur (ie the outcome is missing) up to age of 10 years (if the period of the data collection covers 10 years of age). This may require an additional linkage/search of the cases that have not died with other population sources (see also subsection 3.2 ‘Local analyses’ below).

***2.4. Age at death***

The following definitions and categories (using the WHO and EU definitions[13](#_ENREF_13),[14](#_ENREF_14)) for classifying mortality during the first year of life will be used:

1. early neonatal death (death within the first 0-6 completed days of life)
   1. deaths in first day of life (day 0)
   2. deaths in day 1
   3. deaths in day 2
   4. deaths in day 3
   5. deaths in day 4
   6. deaths in day 5
   7. deaths in day 6
2. late neonatal death (death within the first 7-27 days of life)
3. postneonatal death (death from 28 days to under 1 year)
   1. deaths from 28 days to 2 completed months.
   2. deaths from 3 months to 5 completed months
   3. deaths from 6 months to under 1 year(<365 days)
4. deaths after first year of life (after year 0)
   1. deaths during year 1
   2. deaths during year 2
   3. deaths during year 3
   4. deaths during year 4
   5. deaths during year 5
   6. deaths during year 6
   7. deaths during year 7
   8. deaths during year 8
   9. deaths during year 9

In addition to dates of birth and death, a separate variable on age at death in completed days for the first year of life (year 0) and in completed years for ages 1-9 years up to their 10th birthday (i.e. including 9 years and 364 days) is requested. In particular, for neonatal death (within the first 0-27 days of life), age at death is an important variable, as age could be categorised wrongly if calculated from dates of birth and death in examples when the birth occurred at the end of the previous day and death occurred in the beginning of the next day.

***2.5. Cause of death (see Table 2 “Coding instructions”)***

The EUROlinkCAT Standardisation Committee recommended:

For accidental death use the appropriate “cause of death” codes.

In addition to ICD-9 or ICD-10 codes, please provide any free or standardised text information available in your registry. Please translate the free text information, if not already in English.

All the individual causes of death for each matched child available in the linked file should be used for the aggregated tables to be transferred to the CRR. The local registries/ data providers will run frequencies on causes of death and send these to Ulster University, so that aggregate tables showing common causes of death can be created and included in the analyses syntax scripts.

We do not suggest any major grouping of causes of death at this stage as the classification of causes of death into major groups may differ between the registries.

***2.6. Data on risk factors***

***EUROCAT variables***

Data for the following risk factors are requested (see Table 1 below):

Child’s sex (EUROCAT core variable 4 – SEX),

Plurality (number of babies delivered – singleton or multiple) (EUROCAT core variable 5 – NBRBABY)

Birth weight (grams) (EUROCAT core variable 10 – WEIGHT)

Gestational age (completed weeks) (EUROCAT core variable 11 – GESTLENGTH)

Maternal age (years) (EUROCAT core variable 15 – AGEMO (age of mother at delivery)

Prenatal diagnosis (EUROCAT core variables 19 - WHENDISC (When discovered), 20 – CONDISC (Condition at discovery), 21 – AGEDISC (If prenatally diagnosed, gestational age at discovery), 48-56)

***Socio-economic status (SES):***

Maternal occupation (EUROCAT variable 59 – OCCUPMO (mother’s occupation at time of conception))

Maternal education (EUROCAT variable 91 – MATEDU)

Socioeconomic status of mother (EUROCAT variable 92 – SOCM)

Socioeconomic status of father (EUROCAT variable 93 – SOCF)

Migrant status (EUROCAT variable 94 – MIGRANT)

All the above EUROCAT SES variables have a high percentage of missing data in some participating registries (see Appendix 2).

***Socioeconomic variables among the Mortality variables***

Based on the survey responses from the registries, the following SES variables are available in the mortality databases of the WP3 participants and will be used for this work package (see availability for each registry in Appendix 3):

maternal occupation/employment;

maternal education;

marital status;

non-European Ethnic origin –nationality/ethnicity variables or maternal country of birth/ place of birth/ country of origin as a proxy.

As the SES variables/ proxies for SES are incomplete in the congenital anomaly registries, some registries can link their data to other databases with information on maternal SES. These registries will link three datasets: congenital anomaly data, mortality data and another database with maternal information which will require the maternal ID.

**3. Structure of the local database and local analyses**

***3.1. Structure of the database***

As stated above, the linked data file will be stored securely, either within the local registry or within the organisation doing the linkage. The local database will include all anomaly cases exported from the most recent version of the EDMP and coded according to the EUROCAT Guide 1.4 classification of congenital anomalies. The anomaly cases will include not only the live births involved in the linkage described in section 2 above (both those successfully linked and those unable to be linked), but also all pregnancies resulting in a fetal death from 20 weeks of gestation or TOPFA at any gestational age. All anomaly cases will have the EUROCAT core and additional variables listed in Table 1, including all EUROCAT subgroups for each case, and all linked live births will have in addition the corresponding mortality variables (also listed in Table 1). The procedure of the data transfer to CRR is explained in section 4 below.

**3*.2. Local analyses***

The Ulster University, with support from Queen Mary University of London (QMUL), will produce common syntax scripts which will conform to a common data model to ensure that all variables/proxy variables are standardised across all registries. Registries will use the syntax scripts provided on their linked dataset to generate the tables/ results outlined in the analysis plan. Rules for each registry to generate derived study variables from the existing data and a "dictionary" of every variable in the data with its name, description/definition, coding instructions/ values (in English) will be created.

All cases of congenital anomalies classified into standard EDMP derived EUROCAT subgroups (based on coding in chapter 3.3of theEUROCAT Guide 1.4) will be used: for example but not restricted to, spina bifida, Tetralogy of Fallot, Transposition of great vessels, oro-facial clefts, oesophageal atresia, small intestinal atresia/stenosis, anorectal atresia, diaphragmatic hernia, omphalocele, gastroschisis and certain syndromes.

The specific anomaly subgroups/anomalies for the analyses will be decided after a detailed literature review and discussions with neonatologists and paediatricians, and additional subgroups with very rare anomalies and syndromes will be defined based on ICD9 and ICD10 codes in the corresponding ICD malformation chapters, to ensure that the anomaly subgroups are clinically relevant. In addition, careful power calculations will be performed to ensure that the selected anomaly subgroups have sufficient power to derive meaningful conclusions. All agreed subgroups to use in the analyses will be defined in the common syntax script.

Linked deaths will be stratified by the agreed congenital anomaly subgroups and further by the risk factors listed in section 2.6 above.

Syntax scripts will also be provided to determine the quality of the data to be linked, for example, completeness of each variable, and the completeness of the data linkage. The consistency of the coding and definition of the variables across the registries will be standardised. Standard data checks for unlikely results and outliers will be part of the analysis using the Ulster University common syntax scripts.

Unlinked cases need to be described as they may represent high risk cases, early deaths without a personal ID applied etc, and reported to the Ulster University. A variable indicating if a registry case has been matched with the mortality data, confirmed to be alive or not matched for various reasons has been added to the list of additional variables (MORT\_MATCH, see Tables 1-2). Another addition is a variable quantifying the strength of the linkage (MORT\_MATCH\_STRENGTH) (see the first point in section 5).

As the date of death in mortality data, ‘DEATH\_DATE\_MORT’ variable, is expected to be more complete than the EUROCAT ‘DEATH\_DATE’ variable in a congenital anomaly registry (according to 2008-2012 table, missing data for about 63% on average, ranging from 0% to 99.4%), ‘DEATH\_DATE\_MORT’ variable, will be used for calculation of age at death.

Each registry will run common scripts produced by WP2 (Ulster University) to create standard summary tables and perform centrally-defined analyses.

**4. Data transmission to the Central Results Repository (CRR) and to WP3: Mortality**

The tables and results created by each registry using the supplied syntax scripts will be submitted in Excel or STATA file formats, or other commercially available packages, to the Ulster University via the secure project portal (members’ area on the EUROlinkCAT website). All data submitted will be aggregated - no individual case data will be sent.

Ulster University will then:

1. Compile the data from each participating registry to create the CRR.

2. Generate extracts of the data from the CRR required for studying specific outcomes (i.e. mortality for WP3) of children with congenital anomalies.

3. Submit the relevant CRR extracts to QMUL (Centre for Environmental and Preventive Medicine) to facilitate data validation and provision of analysis/programming support for member registries and the WP3 leaders.

4. Submit the relevant CRR extracts to the WP3 leaders via the secure project portal (members’ area on the EUROlinkCAT website).

The WP3 leaders with their staff will perform pooled analyses based on the data received from the CRR.

**Table 1 Master table of the list of the required variables (in addition to all the subgroups codes provided in EDMP)**

|  |  |  |
| --- | --- | --- |
|  | **EUROCAT core variables plus additional variables with information on linkage, one row of data per case (see also** [**http://www.eurocat-network.eu/aboutus/datacollection/guidelinesforregistration/guide1\_4**](http://www.eurocat-network.eu/aboutus/datacollection/guidelinesforregistration/guide1_4)**)** | |
|  | **EDMP Core variables (shaded in blue)** | |
| **Baby and Mother – Variables 1 to 18** | | |
| 1 | CENTRE | Centre Number |
| 2 | NUMLOC | Local ID of case |
| 3 | BIRTH\_DATE | Date of Birth |
| 4 | SEX | Sex |
| 5\*\* | NBRBABY | Number of babies delivered |
| 6 | SP\_TWIN | Specify twin type of birth, like or unlike, zygosity |
| 7 | NBRMALF | Number of malformed in multiple set |
| 8 | TYPE | Type of birth |
| 9 | CIVREG | Civil registration status |
| 10 | WEIGHT | Birth weight |
| 11 | GESTLENGTH | Length of gestation in completed weeks |
| 12 | SURVIVAL | Survival beyond one week of age |
| 13 | DEATH\_DATE | Date of death |
| 14 | DATEMO | Date of birth of mother |
| 15 | AGEMO | Age of mother at delivery |
| 16\* | BMI | Maternal Body Mass Index |
| 17 | RESIDMO | Mother’s residence code |
| **Diagnosis – Variables 19 to 57** | | |
| 19\*\* | WHENDISC | When discovered |
| 20 | CONDISC | Condition at discovery |
| 21 | AGEDISC | If prenatally diagnosed, gestational age at discovery |
| 24 | KARYO | Karyotype of infant/fetus |
| 25 | SP\_KARYO | Specify karyotype |
| 26\* | GENTEST | Genetic Test |
| 27\* | SP\_GENTEST | Specify genetic test |
| 28 | PM | Post mortem examination |
| 29\*\* | SURGERY | First surgery for malformation performed or planned |
| 30 | SYNDROME | Syndrome |
| 31 | SP\_SYNDROME | Specify Syndrome |
| 32 | MALFO1 | malformation |
| 33 | SP\_MALFO1 | Specify malformation |
| 34 | MALFO2 | As MALFO1 |
| 35 | SP\_MALFO2 | Specify malformation |
| 36 | MALFO3 | As MALFO1 |
| 37 | SP\_MALFO3 | Specify malformation |
| 38 | MALFO4 | As MALFO1 |
| 39 | SP\_MALFO4 | Specify malformation |
| 40 | MALFO5 | As MALFO1 |
| 41 | SP\_MALFO5 | Specify malformation |
| 42 | MALFO6 | As MALFO1 |
| 43 | SP\_MALFO6 | Specify malformation |
| 44 | MALFO7 | As MALFO1 |
| 45 | SP\_MALFO7 | Specify malformation |
| 46 | MALFO8 | As MALFO1 |
| 47 | SP\_MALFO8 | Specify malformation |
| 57# | OMIM | OMIM code / Type of Mendelian Inheritance |
| **Exposure – variables 58 to 78** | | |
| 58\*\* | ASSCONCEPT | Assisted conception (where available) |
| 59## | OCCUPMO | Mother’s occupation at time of conception |
| **Sociodemographic – Variables 91 to 94** | | |
| 91 | MATEDU | Maternal education |
| 92 | SOCM | Socioeconomic status of mother |
| 93 | SOCF | Socioeconomic status of father |
| 94 | MIGRANT | Migrant status |
| **EDMF-derived variables** | |  |
|  | Byear | Year of birth |
|  | birth\_type | Definitions of stillbirths and spontaneous abortions vary between regions. This variable recodes birth type according to EUROCAT’s specifications: cases with gestational age ≥ 20 weeks are re-coded as “stillbirths” (irrespective of the local definition of stillbirth/spontaneous abortion). |
|  | casestatus | Only cases with casestatus = 1 |
|  | al1-al108 | EUROCAT subgroups: (0 = No, 1 = Yes). Based on EUROCAT coding in Guide 1.4 |
|  | mult\_malf | Algorithm for case classification into isolated and multiples |
|  | MM\_Final\_Verdict | This is the final aetiological classification following review of potential multiple cases by a panel of 3 geneticists |
| **Variables in mortality databases to be linked to congenital anomaly cases** | | |
|  | MORT\_MATCH | Match with mortality data or any other population data to establish the vital status (death/alive) |
|  | MORT\_MATCH\_STRENGTH | Strength of match with mortality data |
|  | CHILD\_ID | Child unique ID or other unique identifier |
|  | DATE\_BIRTH | Child’s date of birth |
|  | AGE\_DEATH | Age at death in complete days for the first year (and hours for day 0, where possible) and in complete years after the first year |
|  | AGE\_DEATH\_GR | Categorised age at death |
|  | CHILD\_COUNTRY\_BIRTH | Country of birth of infant |
|  | CHILD\_COUNTRY\_RESIDENCE | Country of residence of infant |
|  | CHILD\_MUNICIPALITY\_RESIDENCE | Municipality of residence of infant |
|  | CHILD\_PROVINCE\_RESIDENCE | Child’s province of residence |
|  | CHILD\_DISTRICT\_RESIDENCE | Child’s district of residence |
|  | CHILD\_COMMUNITY\_LIVING | Child’s community of living |
|  | CHILD\_SEX | Infant’s sex |
|  | GEST\_AGE | Gestational age at birth |
|  | BIRTH\_WEIGHT | Child’s birth weight |
|  | MULT\_BIRTH | Singleton or multiple birth |
|  | REGISTR\_TYPE | Type of birth civil registration of baby |
|  | MOTHER\_REGION\_RESIDENCE | Municipality, Province or Region of residence of the mother to identify non-residents of the registry catchment area |
|  | MOTHER\_DISTRICT\_RESIDENCE | District of mother’s residence |
|  | NON\_EUROPEAN\_ORIGIN | Citizenship/ Nationality of infant |
|  | MAT\_COUNTRY\_BIRTH | Maternal country of birth/ place of birth/ country of origin |
|  | MAR\_STATUS | Marital status |
|  | MAT\_EMPLOY | Mother’s employment status |
|  | MAT\_OCCUP | Maternal occupation |
|  | MAT\_EDUCATION | Maternal education |
|  | MAT\_AGE\_AT\_BIRTH | Maternal age at infant's birth |
|  | DATE\_of\_DEATH | Date of death on death certificate or in mortality database |
|  | CHILD\_COUNTRY\_DEATH | Child’s country of death |
|  | CHILD\_REGION\_DEATH | Region of child’s death |
|  | CHILD\_MUNICIPALITY\_DEATH | Municipality of child’s death |
|  | CHILD\_DISTRICT\_DEATH | District of child’s death |
|  | PLACE\_DEATH | Place of death (home, hospital etc) |
|  | CAUSE\_DEATH\_IA | DIRECT CAUSE OF DEATH (I(a) on death certificate)  Cause of death text -IA |
|  | CAUSE\_DEATH\_1B | Sequence leading to death, i.e. disease or condition leading to 1a - 1(b) on death certificate  Cause of death text -B |
|  | CAUSE\_DEATH\_IC | Disease or condition leading to 1b - 1(c) on death certificate (this could be an underlying cause)  Cause of death text -C |
|  | CAUSE\_DEATH\_ICD9\_3 | Cause of death, first 3 digits ICD9 |
|  | CAUSE\_DEATH\_ICD9\_4th | Cause of death, 4th digit ICD9 |
|  | CAUSE\_DEATH\_II | Contributory cause of death, i.e. conditions contributing to death - (2) on death certificate  Cause of death text -II) |
|  | CAUSE\_DEATH\_TEXT | Cause of death text |
|  | CAUSE\_DEATH\_ICD10\_3 | Cause of death, first 3 digits ICD10 |
|  | CAUSE\_DEATH\_ICD10\_4th | Cause of death, 4th digit ICD10 |
|  | CAUSE\_UNDER\_ICD9 | Underlying cause - ICD9 |
|  | CAUSE\_UNDER\_ICD10 | Underlying cause - ICD10 |
|  | MULT\_CAUSE\_DEATH | Multiple cause of death |

\* New variable In Guide 1.4

\*\* Variable compatible with Guide 1.3, but coding has been extended/modified

# Variable name change only

## Guide 1.4 use ISCO-08 classifications

**Table 2 Coding Instructions for variables in mortality databases** (not given for the EUROCAT variablesas available in the EUROCAT Guide 1.4)

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables in mortality databases to be linked to congenital anomaly cases** | | |  |
|  | **Variable** | **Description** | **Coding** |
|  | MORT\_MATCH | Match with mortality data or other sources | 1=match with mortality data (i.e. death); 2= not death (confirmed to be alive at the time of linkage); 3= not match with mortality data, survival status unknown; 9= not linkable due to insufficient information for linkage |
|  | MORT\_MATCH\_STRENGTH | Strength of match with mortality data | The standardisation of coding will be decided at a later stage depending on the data collected |
|  | CHILD\_ID | Child unique ID or other unique identifier | Personal ID is needed to link to the mortality data |
|  | DATE\_BIRTH | Child’s date of birth | Exact date of birth or month and year of birth if date of birth is unavailable |
|  | AGE\_DEATH | Child’s age at death | Age at death in complete days for the first year (and hours for day 0, where possible) and in complete years after the first year |
|  | AGE\_DEATH\_GR | Group of child’s age at death | 0 = first 24 hrs;  1 = 1 complete day;  2 = 2 complete days;  3 = 3 complete days;  4 – 4 complete days;  5 – 5 complete days;  6 = 6 complete days;  7-= 7-27 days;  8 = 28-365 days;  9= 1 complete year (before 2nd birthday);  10 = 2 complete years (before 3rd birthday);  11 = 3 complete years (before 4th birthday);  12 = 4 complete years (before 5th birthday);  13 = 5 complete years;  14 = 6 complete years;  15 = 7 complete years;  16 = 8 complete years;  17 = 9 complete years (before 10th birthday). |
|  | CHILD\_COUNTRY\_BIRTH | Country of birth of infant |  |
|  | CHILD\_COUNTRY\_RESIDENCE | Country of residence of infant |  |
|  | CHILD\_MUNICIPALITY\_RESIDENCE | Municipality of residence of infant |  |
|  | CHILD\_PROVINCE\_RESIDENCE | Child’s province of residence |  |
|  | CHILD\_DISTRICT\_RESIDENCE | Child’s district of residence |  |
|  | CHILD\_COMMUNITY\_LIVING | Child’s community of living |  |
|  | CHILD\_SEX | Infant’s sex | 1=male; 2=female; 3=indeterminate; 9-= not known |
|  | GEST\_AGE | Gestational age at birth | Gestation age in complete weeks |
|  | BIRTH\_WEIGHT | Child’s birthweight | Birthweight in grams |
|  | MULT\_BIRTH | Singleton or multiple birth | 1=singleton; 2=twin, 3=triplet, 4=quadruplet; 5=quintuplet; 6= sextuplet or more; 7 = Multiple birth, number of babies not known; 8=singleton at time of delivery but known to have been a multiple pregnancy; 9=not known |
|  | REGISTR\_TYPE | Type of birth civil registration of baby | 1=Livebirth  2 = Stillbirth  3 = No civil registration  9 = Not known |
|  | MOTHER\_REGION\_RESIDENCE | Municipality, Province or Region of residence of the mother to identify non-residents of the registry catchment area |  |
|  | MOTHER\_DISTRICT\_RESIDENCE | District of mother’s residence |  |
|  | NON\_EUROPEAN\_ORIGIN | Citizenship/ Nationality of infant |  |
|  | MAT\_COUNTRY\_BIRTH | Maternal country of birth/ place of birth/ country of origin | To use as a proxy for the child’s ethnicity |
|  | MAR\_STATUS | Marital status | 0=Not stated; 1=Single; 2=Married 3=Widow; 4=Divorced |
|  | MAT\_EMPLOY | Mother’s employment status |  |
|  | MAT\_OCCUP | Maternal occupation |  |
|  | MAT\_EDUCATION | Maternal education |  |
|  | MAT\_AGE\_AT\_BIRTH | Maternal age at infant's birth |  |
|  | DATE\_of\_DEATH | Date of death on death certificate or in mortality database | Exact date of death or month and year of birth if date of death is unavailable |
|  | CHILD\_COUNTRY\_DEATH | Child’s country of death |  |
|  | CHILD\_REGION\_DEATH | Region of child’s death |  |
|  | CHILD\_MUNICIPALITY\_DEATH | Municipality of child’s death |  |
|  | CHILD\_DISTRICT\_DEATH | District of child’s death |  |
|  | PLACE\_DEATH | Place of death | 1 = home  2 = hospital  3 =ambulance  4 = school  5 =other(specify)  9 = Not known |
|  | CAUSE\_DEATH\_IA | DIRECT CAUSE OF DEATH (I(a) on death certificate)  *Disease or condition leading directly to deat*h | Text together with the code should be given according to the International Classification of Diseases, Tenth or Ninth Revision, Clinical Modification) |
|  | CAUSE\_DEATH\_IB | Sequence leading to death, i.e. disease or condition (I(b) on death certificate)  1 (b) *Other disease or condition, if any, leading to* I(a) | Text together with the code should be given according to the International Classification of Diseases, Tenth or Ninth Revision, Clinical Modification |
|  | CAUSE\_DEATH\_IC | I(c) *other disease or condition, if any, leading to* I(b) | Text together with the code should be given according to the International Classification of Diseases, Tenth or Ninth Revision, Clinical Modification |
|  | CAUSE\_DEATH\_d |  | Text together with the code should be given according to the International Classification of Diseases, Tenth or Ninth Revision, Clinical Modification |
|  | CAUSE\_DEATH\_ICD9\_3 | Cause of death, first 3 digits ICD9 |  |
|  | CAUSE\_DEATH\_ICD9\_4th | Cause of death, 4th digit ICD9 |  |
|  | CAUSE\_DEATH\_II | Contributory cause of death, i.e. conditions contributing to death - (2) on death certificate  Cause of death text -II |  |
|  | CAUSE\_DEATH\_TEXT | Cause of death text |  |
|  | CAUSE\_DEATH\_ICD10\_3 | Cause of death, first 3 digits ICD10 |  |
|  | CAUSE\_DEATH\_ICD10\_4th | Cause of death, 4th digit ICD10 |  |
|  | CAUSE\_UNDER\_ICD9 | Underlying cause of death - ICD9  I (d) *other disease or condition, if any, leading to* I(c) |  |
|  | CAUSE\_UNDER\_ICD10 | Underlying cause of death - ICD10 |  |
|  | CAUSE\_DEATH\_II | Contributory causes, i.e. other significant conditions contributing to death (2 on death certificate)  2 *Other significant conditions* **Contributing to death** *but no*t *related to the disease o*r *condition causing i*t | Text together with the code should be given according to the International Classification of Diseases, Tenth or Ninth Revision, Clinical Modification |
|  | MULT\_CAUSE\_DEATH | Multiple cause of death |  |

**5. Statistical analysis**

The statistical analysis will be performed in three stages.

1) A series of data quality checks specified centrally by Ulster University will be performed by each participating registry to ensure accurate linkage has occurred and to provide measures of the strength of the linkage. Registries will be asked to produce frequency tables on the linked variables (including proportions not linked, missing) to assess data quality. Data quality criteria for inclusion in full analysis will be agreed, and the registries that do not meet these criteria may need to be excluded.

2) Participating registries will run a first stage of pre-specified analyses using individual-based data. This will include cross-tabulations for outcomes of interest (for example, infant survival and survival >1 year by prenatal diagnosis, survival by congenital anomaly subgroup, survival by maternal age). For relatively common congenital anomalies, logistic regression or survival analysis will be used in each registry to analyse the contribution of various risk factors to the risk of mortality at different child’s ages. The odds of dying will be explored among children with a specific congenital anomaly; crude and adjusted odds ratios for pre-specified confounders (e.g. maternal age, SES) will be calculated. These analyses will be run separately by each registry and the results submitted to the CRR using a secure data transfer procedure and data suppression if required.

3) In the second stage of the analysis, the WP3 team will analyse the data submitted to the CRR (aggregated tables and individual analytical results from the independent standardised databases). This approach was used in a previous similar EUROCAT study (EUROmediCAT).[16](#_ENREF_16) The WP3 team will combine estimates produced by each individual registry in a meta-analysis using a random effects model. For example, we will estimate the effect of a prenatal diagnosis of a congenital anomaly on risk of death at > 1 year in random effects meta-analysis, by producing combined unadjusted and adjusted (for registry, year of diagnosis, maternal SES) odds ratios (ORs) for the multi-registry effects estimates. We will also assess the between-registry heterogeneity in these associations.

Geographical differences will be evaluated by the analysis of indicators calculated by each registry. Both crude and adjusted ORs for risk factors will be used for this aim. In order to contextualise and better interpret the results, as no controls are required, general information on mortality (e.g. neonatal mortality) and prevalence of the risk factors (e.g. birth weight, maternal age, etc.) in the general population covered by the registry will be collected. Furthermore, the data on TOPFA and fetal deaths, including information on time of prenatal diagnosis, will be analysed.

Table 3a below shows the estimated number of live births and deaths up to 10 years of age for the participating registries.

**Table 3a.** Estimated number of live births and deaths up to 10 years of age for 21 EUROCAT registries

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Congenital Anomaly Registry | Start year | Estimated total births in population to 2014 | Estimated live births with an anomaly up to 2014 | Estimated deaths\* |
| Belgium: Antwerp | 1995 | 372,394 | 8083 | 1059 |
| Croatia: Zagreb | 1995 | 136,979 | 2232 | 292 |
| Denmark: Odense | 1995 | 106,026 | 2418 | 317 |
| Finland | 1995 | 1179,314 | 44869 | 5878 |
| France: Ile de la Reunion | 2002 | 189,647 | 3855 | 505 |
| France: Paris | 1997 | 598,208 | 13335 | 1747 |
| Germany: Saxony-Anhalt | 1995 | 308,747 | 8821 | 1156 |
| Italy: Emilia Romagna | 1995 | 674,044 | 11447 | 1500 |
| Italy: Tuscany | 1995 | 565,131 | 9827 | 1287 |
| Malta | 1995 | 847,69 | 2470 | 324 |
| Netherlands: Northern | 1995 | 373,474 | 8567 | 1122 |
| Norway | 1999 | 836,535 | 24255 | 3177 |
| Spain: Basque | 1995 | 318,788 | 4883 | 640 |
| Spain: Valencia | 2007 | 409,296 | 7438 | 974 |
| UK: East Midlands | 1998 | 1,151,533 | 18549 | 2430 |
| UK: North | 2000 | 484,393 | 8617 | 1129 |
| UK: South West | 2005 | 500,374 | 11671 | 1529 |
| UK: Thames Valley | 1995 | 362,051 | 5142 | 674 |
| UK: Wales | 1998 | 572,558 | 18239 | 2389 |
| UK: Wessex | 1995 | 561,192 | 7771 | 1018 |
| Ukraine: West | 2005 | 306,980 | 6166 | 808 |
| **Total** |  | 10,092,433 | 228,655 | 29,954 |

\*Deaths were estimated using 10-year survival (86.9%) estimates from the North of England study (Tennant et al, 2010)[10](#_ENREF_10)

**Table 3b.** Estimated number of live births with deaths up to 10 years of age for anomaly subgroups listed on page 6 of the protocol for 21 EUROlinkCAT registries, 1995-2014

|  |  |  |  |
| --- | --- | --- | --- |
| Congenital Anomaly subgroup | Estimated live births with an anomaly | 10-year mortality (%) | Estimated deaths\* |
| Spina bifida | 1732 | 31.3 | 542 |
| Tetralogy of Fallot | 3035 | 16.9 | 513 |
| Transposition of great vessels | 3239 | 19.7 | 638 |
| Oro-facial clefts | 14784 | 2.4 | 355 |
| Oesophageal atresia | 2470 | 6.7 | 165 |
| Duodenal atresia/stenosis | 1362 | 4.6 | 63 |
| Intestinal atresia/stenosis, other parts of small intestine† | 963 | 7.7 | 74 |
| Anorectal atresia | 2569 | 1.2 | 31 |
| Diaphragmatic hernia | 2181 | 42.9 | 936 |
| Omphalocele | 1310 | 12.8 | 168 |
| Gastroschisis | 2398 | 6.3 | 151 |
| Skeletal dysplasias | 1088 | 34.7 | 378 |
| Cranyosynostosis | 2902 | 2.4 | 70 |

\*Deaths were estimated using anomaly-specific 10-year survival estimates from the North of England study (Tennant et al, 2010)[10](#_ENREF_10)

† 5—year survival was used for estimation as 10-year survival not provided

Table 3b shows that for a number of congenital anomaly subgroups, there will be sufficient number of deaths to do the adjusted analyses. For subgroups with few cases, unadjusted analyses will be performed. The results for these subgroups can be compared across the registries with differing TOPFA rates.

For those registries contributing to the WP3 Mortality and WP4 Morbidity studies, this data request will be linked to the protocol and data request of WP4 Morbidity as they are interrelated in terms of study outcomes and the variables that affect both mortality and morbidity outcomes. The factors, such as type and timing of surgery, if applicable, history of infections, respiratory diseases, admissions and duration of hospital stay, characterise severity of the condition and influence children’s survival. The influence of these factors will also be investigated in the survival analyses.

**6. Data archiving and destruction**

1. All results generated from EUROlinkCAT will be archived at the Ulster University. These will include the pooled analyses from each WP.
2. Destruction of CRR data will occur 20 years after the completion of the EUROlinkCAT project, at which point it is believed such data will no longer be of use.
3. Each WP leader/institution is responsible for ensuring the destruction of any data five years after the completion of the EUROlinkCAT project.
4. Any duplicate datasets held at QMUL will be destroyed five years after the completion of the EUROlinkCAT project.

**References**

1. Wolfe I, Thompson M, Gill P, Tamburlini G, Blair M, van den Bruel A, Ehrich J, et al. Health services for children in western Europe. *Lancet* 2013;381:1224-1234.

2. Rosano A, Botto LD, Botting B, Mastroiacovo P. Infant mortality and congenital anomalies from 1950 to 1994: an international perspective. *J Epidemiol Community Health* 2000;54:660-666.

3. Office for National Statistics. Childhood, Infant and Perinatal Mortality in England and Wales: 2012. 2014. Accessed at: 15 May 2017 from: https://[www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/childhoodinfantandperinatalmortalityinenglandandwales/2014-01-30](http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/childhoodinfantandperinatalmortalityinenglandandwales/2014-01-30).

4. Rankin J, Tennant PW, Bythell M, Pearce MS. Predictors of survival in children born with Down syndrome: a registry-based study. *Pediatrics* 2012;129:e1373-1381.

5. Wu J, Morris JK. The population prevalence of Down's syndrome in England and Wales in 2011. *Eur J Hum Genet* 2013;21:1016-1019.

6. Shin M, Kucik JE, Siffel C, Lu C, Shaw GM, Canfield MA, Correa A. Improved survival among children with spina bifida in the United States. *J Pediatr* 2012;161:1132-1137.

7. Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Erickson JD. Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979-1997. *Circulation* 2001;103:2376-2381.

8. Gordon BM, Rodriguez S, Lee M, Chang RK. Decreasing number of deaths of infants with hypoplastic left heart syndrome. *J Pediatr* 2008;153:354-358.

9. Erikssen G, Liestol K, Seem E, Birkeland S, Saatvedt KJ, Hoel TN, Dohlen G, et al. Achievements in congenital heart defect surgery: a prospective, 40-year study of 7038 patients. *Circulation* 2015;131:337-346; discussion 346.

10. Tennant PW, Pearce MS, Bythell M, Rankin J. 20-year survival of children born with congenital anomalies: a population-based study. *Lancet* 2010;375:649-656.

11. Copeland GE, Kirby RS. Using birth defects registry data to evaluate infant and childhood mortality associated with birth defects: an alternative to traditional mortality assessment using underlying cause of death statistics. *Birth Defects Res A Clin Mol Teratol* 2007;79:792-797.

12. Boyd PA, Haeusler M, Barisic I, Loane M, Garne E, Dolk H. Paper 1: The EUROCAT network--organization and processes. *Birth Defects Res A Clin Mol Teratol* 2011;91 Suppl 1:S2-15.

13. The European Commission. COMMISSION REGULATION (EU) No 328/2011 of 5 April 2011. *Official Journal of the European Union* 2011;54:22-24.

14. World Health Organization. International statistical classification of diseases and related health problems, 10th revision. Fifth Edition. Volume 2, Instruction manual. 2016. <http://apps.who.int/classifications/icd10/browse/Content/statichtml/ICD10Volume2_en_2016.pdf?ua=1&ua=1>.

15. Dattani N, Rowan S, Office for National Statistics. Causes of neonatal deaths and stillbirths: a new hierarchical classification in ICD-10. *Health Stat Q* 2002;15:16-22.

16. Garne E, Vinkel Hansen A, Morris J, Jordan S, Klungsoyr K, Engeland A, Tucker D, et al. Risk of congenital anomalies after exposure to asthma medication in the first trimester of pregnancy - a cohort linkage study. *BJOG* 2016;123:1609-1618.

**Appendix 1**

**Sources of Mortality data for Linkage (as reported by the EUROCAT registries)**

* Antwerp: the Flemish Agency for Care and Health
* Basque Country: Health System Database or more specific mortality database for Spain which is the “Indice Nacional de defunciones”.
* BINOCAR (England): NHS Digital will provide mortality data from the Office for National Statistics-

<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/qmis/childmortalitystatisticsqmi>

* + - * Denmark: Statistics Denmark
* Emilia Romagna: data from three databases will need to be linked as deaths in 1st year are treated completely differently. ISTAT database, a national mortality system.
* Finland: THL register linked with Statistics Finland data on deaths.
  + - * Malta: the Mortality Register
      * N Netherlands: Statistics Netherlands (Centraal Bureau voor de Statistiek(CBS) -<http://statline.cbs.nl)>
      * Norway: the Medical Birth Registry of Norway and the Cause of Death Registry
      * Paris: INSEE (National Institute of Statistics and Economic Studies) to identify the national register for the identification of individuals (RNIPP - National Directory for the Identification of Natural Persons) and CépiDc (<http://www.cepidc.inserm.fr/>)- service in charge of the national statistics of causes of death, incorporated in INSERM, national institution of medical research.
      * Reunion: Data base of the hospitals (department of medical information via CROSSWAY), INSEE La Réunion, Mayotte (Service of Studies & Diffusion), and local database in our services in link with the registry
      * Saxony-Anhalt: Linkage with regional death registration "Statistisches Landesamt Sachsen-Anhalt"
      * Tuscany: Regional Mortality Registry, Regional inhabitant Registry, Report for National Institute of Statistics.
      * Ukraine: Regional Children Hospital Statistics.
      * Valencia: Mortality Registry.
      * Wales: ONS Mortality.
      * Zagreb: Croatian Institute of Public Health.

**Appendix 2**

**The variation in the percentage of missing (unknown, invalid) data between the EUROCAT congenital anomaly registries for 2008-2012**:

**OCCUPMO** (mother’s occupation at time of conception)**:** the percentage of missing data ranges from the lowest of 0.9% in Ukraine and 7.7% in Reunion to the highest of 94.5% in the Basque Country, 96.3% in BINOCAR and 100% in Norway based on 2008-2012 data.

**MATEDU** (maternal education)**:** the percentage of missing data ranges from the lowest of 0.5% in Ukraine and 4.6% in Zagreb to the highest of 100% in Paris, BINOCAR, Norway, Saxony Anhalt and Basque Country.

**SOCM** (socioeconomic status of mother)**:** the percentage of missing data ranges from the lowest of 4.0% in Ukraine and Zagreb to the highest of 100% in Antwerp, Basque Country, N Netherlands, Norway, Paris, Tuscany and Saxony Anhalt.

**SOCF** (socioeconomic status of father)**:** the percentage of missing data ranges from the lowest of 7.4% in Ukraine to the highest of 100% in Antwerp, Basque Country, BINOCAR, N Netherlands, Norway, Paris, Tuscany and Saxony Anhalt.

**MIGRANT (m**igrant status)**:** the percentage of missing data ranges from the lowest of 0.0% in Malta and 0.1% in Ukraine to the highest of 100% in Antwerp, BINOCAR, N Netherlands, Norway, Paris and Tuscany.

**Appendix 3 – Availability of the Mortality variables -** For each registry, a registry-specific table has been attached separately.