

**EUROlinkCAT WP4 - Risk factors for morbidity in congenital anomaly affected children up to ten years of age (Geographical variation)**

# Statistical analysis plan (version 3)

**Aim:** to determine the impact of risk factors that might explain geographical differences in morbidity among congenital anomaly (CA) affected children.

## Scope of the risk factors study in relation to the WP4 long-term morbidity study

The WP4 long-term morbidity study will assess length of hospital stay and other health outcomes in children with CAs and assess if there is geographical variation in these outcomes across Europe. This study will assess if risk factors contribute to any differences in health outcomes across Europe.

## Case and control definition

Cases are children with CAs as defined by the EUROlinkCAT subgroup list. The CA subgroups included in this analysis are based on those used in the WP4 part 1 study:

Unless specified otherwise the CA groups used in the below tables are:

* 1. All anomalies subgroup (al1)
	2. Isolated anomalies[[1]](#footnote-1) (mult\_malf=A,R,N,I) (spina bifida, hydrocephalus, severe microcephaly, CHD, severe CHD, transposition of the great vessels, VSD, ASD, Tetralogy of Fallot, coarctation of the aorta, PDA, cleft lip with or without cleft palate, cleft palate, oesophageal atresia, anorectal atresia, diaphragmatic hernia, gastroschisis, multicystic renal dysplasia, congenital hydronephrosis, hypospadias, limb reduction defects, club foot, hip dislocation and craniosynostosis)
	3. Chromosomal anomalies – Down syndrome (all, with CHD, without CHD), Turner syndrome
	4. Rare anomalies - anomalies of the Corpus Callosum and Di George syndrome.

Controls are children without CAs i.e. any child diagnosed with a Q-code (ICD10) or 740-759 code (ICD9) and/or Di George syndrome at any time in the health care databases will be excluded. The minimum gestational age for controls is 24 weeks.

Where possible, controls will be all children in the population born in the same geographical area and within the same time period that were not recorded in the EUROCAT registry. Where necessary a random 10% sample of controls will be used. Control children are not available for Ile de la Reunion.

**Note:** Due to the number of outcomes and risk factors, there are NUMEROUS possible analyses. **This analysis will therefore be restricted to the health outcomes that showed significant geographical heterogeneity as reported in the long-term morbidity study.** **Similarly, if numbers are small crude analyses alone will be conducted.** Potential risk factors with ≥ 20% missing information in a registry means the registry will be excluded from analysis of that specific risk factor.

There are three types of outcome measure

* 1. Continuous
		1. Length of stay
		2. Length of intensive care stay
	2. Binary
		1. Ever admitted hospital
		2. Ever been in hospital for > 10 days
		3. Ever admitted intensive care
		4. Ever had ventilation
	3. Rare Binary
		1. Ever diagnosed cancer, CP etc - might be very rare cases/controls

## Risk Factors / Confounders

There are three types of risk factors / confounders

* 1. Cohort / Birth Year (i.e. time)
	2. Risk factors contributing to the risk of CA and morbidity
		1. Gender
		2. Multiple birth
		3. Maternal age
		4. Non-resident/migrant status
		5. SES
	3. Risk factors for morbidity affected by CA
		1. Gestational Age
		2. Birth weight

**Notes on risk factors**

Gender will not be examined as a risk factor in hypospadias.

Cases classified as non-national, but exact nationality unknown will be excluded from analysis of non-European ethnic origin. If numbers are too small, then analysis will be based on national versus non-national only, although this will not be able to address non-European ethnic origin as a potential risk factor)

Maternal SES – maternal education, deprivation and occupation have been combined into a proxy SES variable.

Birth weight Is not being investigated as it is so correlated with gestational age.

Maternal BMI is poorly completed and is not being investigated as a risk factor.

**Note:** some registries may not have data for children aged 5-9 years. For these registries we would just examine the risk for <1 year, 1-4 years or 0-4 years only.

# The following tables are based on the Tables for the WP4 Part 1 study (LONG-TERM MORBIDITY), dated 17/06/2019, led by ester.

# Denominators

### Table 1a-c Number of live births with data available at each age (exclude deaths in age group) across risk factors

Record the distribution of risk factors among cases and controls alive at the start of each age group: total numbers by time period, gender, multiple birth, gestational age category, maternal age group, maternal non-resident/migrant status and SES. For each registry, inclusion of any risk factor in analysis depends on the registry data quality reports).

# Hospital admission and Length of hospital stay

### Table 2a-g Hospitalisation and length of stay in controls and CA subgroups by gender among those < 1 year, 1-4 years and 5-9 years.

## Hospital admission example questions:

## Does maternal age increase the risk of admission?

## If maternal age does increase the risk of admission doe this explain any country differences?

## Does the effect of maternal age differ between countries?

For those 0-9 years

1. Each registry will conduct a Poisson regression analysis to examine the Incidence Rate Ratio (IRR) of admission in cases vs controls with the risk factor as a covariate.

### For those <1, 1-4 and 5-9 years old the number and Kaplan-Meier % of cases (across groups specified in WP4 part 1) and controls with ≥0.5 days in hospital will be determined across risk factors.

1. Each registry will conduct a Poisson regression analysis to examine the Incidence Rate Ratio (IRR) of admission in cases vs controls with the risk factor as a covariate.
2. Separate Poisson Regression analyses will also be done for each level of the risk factor producing for example 2 IRRs for gender (male and female).

The data from all registries will be combined in meta-analysis. Combining the Poisson regression coefficients obtained in a) will determine if the risk factor increases the risk of admission in those 0-9 years old. The I2 statistic (heterogeneity) will indicate if the effect of for example maternal age differs across registries.

Combining the Poisson regression coefficients obtained in b) will determine if the risk factor increases the risk of admission in a particular age group. The I2 statistic (heterogeneity) will indicate if the effect of for example maternal age differs across registries.

Multilevel meta-analysis will be conducted using the coefficients produced in c) to compare the risk of admission among the same level of the risk factor across registries for example among males < 1 year old across all registries. Any remaining variation due to country differences will be adjusted for the level of the risk factor.

Depending on numbers it may only be possible to conduct the proposed a) or a/b) analyses.

Length of hospital stay example questions:

## Does maternal age increase LOS?

## If maternal age does increase LOS does this explain any country differences?

## Does the effect of maternal age on LOS differ between countries?

For those 0-9 years

1. Each registry will conduct negative binomial regression analysis to determine the IRR for LOS in cases compared to controls with the risk factor as a covariate.

For those <1, 1-4 and 5-9 years old the median hospital stay (95% CI) for cases and controls with ≥0.5 days in hospital will be determined across risk factors.

1. Each registry will conduct negative binomial regression analysis to determine the IRR for LOS in cases compared to controls with the risk factor as a covariate.
2. Separate negative binomial regression analyses will also be done for each level of the risk factor producing for example 2 IRRs for gender (male and female).

Depending on numbers it may only be possible to conduct the proposed d) or d/e) analyses.

The data from all registries will be combined in meta-analysis. Combining the Poisson regression coefficients obtained in d) will determine if the risk factor increases the LOS. The I2 statistic (heterogeneity) will indicate if the effect of for example maternal age differs across registries.

Combining the Poisson regression coefficients obtained in e) will determine if the risk factor increases the LOS in a particular age group. The I2 statistic (heterogeneity) will indicate if the effect of for example maternal age differs across registries.

Multilevel meta-analysis will be conducted using the coefficients produced in f) to compare the increased LOS among the same level of the risk factor across registries for example among males < 1 year old across all registries. Any remaining variation due to country differences will be adjusted for the level of the risk factor.

### Table 3 Results of regression analysing the effect of the risk factors on risk of admission to hospital, and length of stay, in all CA cases compared to controls at <1, 1-4, 5-9 and 0-9 years of age

Adjusted regression analysis will be conducted for the all anomalies group only to determine the effect of the risk factors on risk of admission to hospital, and length of stay, in CA cases compared to controls at <1, 1-4, 5-9 and 0-9 years of age. The regression coefficients obtained will be combined in a meta-analysis.

# Ever been in hospital > 10 days, intensive care admission and Ventilation use

## Example questions:

## Does maternal age increase the risk of being in hospital for > 10 days?

## If maternal age increases the risk of intensive care admission doe this explain any country differences?

## If maternal age increases the risk of ventilator use does the effect of maternal age differ between countries?

# Table 4 Results of Poisson regression analysing the effect of the risk factors on risk of ever having been in hospital > 10 days, intensive care admission and ventilation use, in all CA cases compared to controls at <1, 1-4, 5-9 and 0-9 years of age. Adjusted regression analysis may need to be restricted to the 0-9 year age group. The regression coefficients obtained will be combined in a meta-analysis.

# Length of intensive care stay

# Tables have not been drafted for length of intensive case stay as these data are not available in all registries. Depending on the results of the first WP4 study and the numbers of cases and controls with this data available a limited exploration may be possible for all anomalies (using the same approach as for LOS in hospital).

# Diagnosis of cancer, cerebral palsy, seizures/epilepsy, renal failure, hearing loss, vision impairments and blindness, "any injury/poisoning" and battered child.

### *Example questions:*

### *Does maternal age increase the risk of developing cancer?*

### *If maternal age does increase the risk of developing cancer does this explain any country differences?*

### *Does the effect of maternal age on the risk of developing cancer differ between countries?*

### Table 5A-C Risk factors for select co-morbidities <1, 1-4 and 5-9 (all anomalies) (see Morbidity Part 1 Sheet 9A)

Poisson regression will be used to determine if at <1, 1-4 and 5-9 years of age there is a difference between cases (all CA cases together – no subgroup analysis)and controls in terms of co-morbidities (various) and if this is related to time period, sex, multiple birth, gestational age, maternal age, maternal non-resident/migrant status, and SES.

Depending on the available numbers it may be necessary to combine these analyses into a single analysis looking at the 0-9-year age group.

### Table 6A-E Risk factors for select co-morbidities 0-9 years of age selected subgroups (see Morbidity Part 1 Sheet 9B)

Poisson regression will be used to determine if there is a difference between cases (select subgroups only – as per WP4 part 1 study)and controls in terms of co-morbidities (cancer, cerebral palsy, seizures/epilepsy, renal failure and visual impairments and blindness) up to the age of 10 (0-9 years) and if this is related to time period, sex, multiple birth, gestational age, maternal age, maternal non-resident/migrant status and SES.

The regression coefficients obtained in Tables 3 and 4 will be combined in a meta-analysis.

1. Anomalies selected based on a livebirth prevalence of ≥1.75 per 10,000 births [↑](#footnote-ref-1)