**Algorithms for use of health care databases in the surveillance of congenital anomalies to improve the quality of the data**

Protocol for a EUROlinkCAT WP6 study

**Aim of this study:**

To develop algorithms for use of health care databases (HCD) in the surveillance of congenital anomalies to improve the quality of the data

**Milestone 23, UNIFE, Month 39 (November 2019):**

Protocol for analysis plan prepared for Algorithm for using congenital anomaly data from hospital discharge databases, available on membership-only section of website

**Deliverable 13/D6.1, UMCG, Month 54 (June 2021):**

Report on the evaluation of specific congenital anomaly coding in health care databases including a computer algorithm to improve these codes

**Institution responsible for algorithm:**

UNIFE Università degli Studi di Ferrara (Dr Gianni Astolfi, Dr Amanda J Neville)

**Other partners: from grant agreement**

1 SGUL UK (J Morris)

2 RSD Denmark (E Garne)

5 UNIFE Emilia Romagna, Italy (A Astolfi, A J Neville)

8 UMCG Northern Netherlands (H de Walle)

All EUROCAT registries will be included in the questionnaire stage

**Background**

Electronic healthcare databases can be a source of useful data for congenital anomaly surveillance and research, but only after careful evaluation. Databases often have different aims such as financial reporting and use another coding system than used in the anomaly surveillance network. Whereas congenital anomaly registries in the beginning depended on individual case identification – either sent to the registry by clinicians or actively searched for in clinical records, the last decade has seen an increase in use of electronic healthcare records as data source for individual cases, and the possibility for registries to ascertain cases by searching for congenital anomaly codes in hospital discharge database.

Registries and researchers using databases for studying congenital anomalies may use algorithms to discriminate between true cases and suspected or minor cases (Astolfi et al 2016). As there are many grey zones in the definition of major congenital anomalies, access to medical records including results of specific examinations (MR scan, echocardiography, genetic tests, post-mortem examinations) may still be necessary for correct interpretation of the cases (Tairou et al 2006). An example for this is atrial septal defect (ASD) where an echocardiography performed in the neonatal period in most cases will show a flow over the atrial septum, as the foramen ovale from fetal life has not yet closed. Many clinicians will code this as ASD in the discharge letter despite the benign nature of this finding (Garne et al 2012).

**Limitations of heath care databases for CA registration**

1. Limited number of diagnosis can be recorded (e.g. SDO Italy only 6)
2. Limited type of diagnosis codes can be recorded
3. No written text description is available
4. Frequent use of generic codes (e.g. Other anomaly of face and neck)
5. Codes for reimbursement so more serious condition recorded.
6. Where no care is required CA are missed (e.g. Downs syndrome Sicily)
7. High number of minor anomalies recorded
8. Conditions that may resolve after birth recorded (e.g. PDA at <37 weeks GA)
9. The database originated mainly for administrative and not epidemiological-scientific aims.
10. Detail in the clinical description of the defect rarely reaches the fifth digit of the ICD9-CM code.
11. Inaccurate identification and classification of anomalies. Some anomalies have an ICD9-CM macrocode that includes several ICD9-BPA or ICD10 codes
12. Inappropriate codings (wrong ICD9, ICD9 of adult and not newborn e.g. ICD10 hydronephrosis)
13. Generic codings such as "Not otherwise specified" (NAS), "Without other indications" (SAI) and "Not indicated elsewhere" (NIA) which make the case doubtful and of poor reliability
14. Regression of malformation during the first months of life. Some CA regress over time (heart and kidney defects) requiring further confirmation and follow-up after the neonatal period.
15. The confirmation of the case requires medical evaluation and the need to request clinical information available in clinical file, with considerable expenditure of time and energy

**Action Plan**

The plan of work to develop an algorithm that can be used by EUROCAT registries to obtain high quality data from health care databases (HCD) will be developed *ex novo* using the experience of the EUROCAT registries in data extraction The algorithm logic will follow EUROCAT Guide 1.4 <https://eu-rd-platform.jrc.ec.europa.eu/eurocat/data-collection/guidelines-for-data-registration#inline-nav-2> and could be applied to any registry. The results will be open access.

1. Send a questionnaire to registries to see if they use HCD how they use it and if they have any algorithms/rules of their own. The questionnaire will be developed using the list of limitations of HCD listed above.
2. Compare the algorithms to other algorithms developed within EUROCAT (including E Garne’s algorithm for CHD, the multiple malformations algorithm)
3. Use data from WP6 tables to suggest any alterations / differences between countries
4. Suggest any additional tables to be produced (for example diagnosis having to occur more than once – table looking at numbers if this rule was applied and if it was not applied)
5. Produce revised detailed algorithm
6. Publish paper with the detailed algorithm that can be used and programmed by others in their own computer languages.

The work of Task 6.2 will be assisted by the TASK 6.1 findings regarding coding.

**Publication of results**

The study will be published in a high-impact peer-review journal with open access and with authorship according to EUROlinkCAT criteria. Submission expected by month 55 (July 2021)

The results will also be included in deliverable D8.3: “Report to EU institutions hosting health care databases with guidelines for improving the quality of the congenital anomaly coding” by month 57 (September 2021)

**References:**

[Astolfi G](https://www.ncbi.nlm.nih.gov/pubmed/?term=Astolfi%20G%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Ricci P](https://www.ncbi.nlm.nih.gov/pubmed/?term=Ricci%20P%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Calzolari E](https://www.ncbi.nlm.nih.gov/pubmed/?term=Calzolari%20E%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Neville A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Neville%20A%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Pironi V](https://www.ncbi.nlm.nih.gov/pubmed/?term=Pironi%20V%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Santoro M](https://www.ncbi.nlm.nih.gov/pubmed/?term=Santoro%20M%5BAuthor%5D&cauthor=true&cauthor_uid=27290890), [Bianchi F](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bianchi%20F%5BAuthor%5D&cauthor=true&cauthor_uid=27290890). [[Validation of an algorithm for identifying cases with congenital malformations by using hospital discharge records].](https://www.ncbi.nlm.nih.gov/pubmed/27290890) Epidemiol Prev. 2016 Mar-Apr;40(2):124-30. Italian.

# [Tairou F](https://www.ncbi.nlm.nih.gov/pubmed/?term=Tairou%20F%5BAuthor%5D&cauthor=true&cauthor_uid=17306063), [De Wals P](https://www.ncbi.nlm.nih.gov/pubmed/?term=De%20Wals%20P%5BAuthor%5D&cauthor=true&cauthor_uid=17306063), [Bastide A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bastide%20A%5BAuthor%5D&cauthor=true&cauthor_uid=17306063). Validity of death and stillbirth certificates and hospital discharge summaries for the identification of neural tube defects in Quebec City. [Chronic Dis Can.](https://www.ncbi.nlm.nih.gov/pubmed/?term=terminations+of+pregnancy+and+congenital+anomaly+and+hospital+discharge+database) 2006;27(3):120-4.

Garne E, Olsen MS, Johnsen SP, Hjortdal V, Andersen HØ, Nissen H, Søndergaard L, Videbaek J; Danish Register of Congenital Heart Disease. [How do we define congenital heart defects for scientific studies?](https://www.ncbi.nlm.nih.gov/pubmed/22010848) Congenit Heart Dis. 2012 Jan-Feb;7(1):46-9.

Rachel A. Charlton, John G. Weil, Marianne C. Cunnington and Corinne S. de Vries. Identifying Major Congenital Malformations in the UK General Practice Research Database (GPRD). Drug Saf 2010; 33 (9): 741-750

Charlton RA, Neville AJ, Jordan S, Pierini A, Damase-Michel C, Klungsøyr K, Andersen AM, Hansen AV, Gini R, Bos JH, Puccini A, Hurault-Delarue C, Brooks CJ, de Jong-van den Berg LT, de Vries CS. [Healthcare databases in Europe for studying medicine use and safety during pregnancy.](https://www.ncbi.nlm.nih.gov/pubmed/24664855) Pharmacoepidemiol Drug Saf. 2014 Jun;23(6):586-94

Garne E, Dolk H, Loane M, Wellesley D, Barisic I, Calzolari E, Densem J; EUROCAT Working Group. Paper 5: Surveillance of multiple congenital anomalies: implementation of a computer algorithm in European registers for classification of cases. Birth Defects Res A Clin Mol Teratol. 2011 Mar;91 Suppl 1:S44-50.

Astolfi G, Bianchi F, Lupi C, Napoli N, Neville A, Verdini E, Verzola A, Calzolari E. [Using hospital discharge records, birth certificates and a birth defects registry for epidemiological and public health purposes: experience in Emilia-Romagna region (northern Italy)]. Epidemiol Prev. 2013 Jul-Oct;37(4-5):279-88.

Canova C, Simonato L, Barbiellini Amidei C, Baldi I, Dalla Zuanna T, Gregori D, Danieli S, Buja A, Lorenzoni G, Pitter G, Costa G, Gnavi R, Corrao G, Rea F, Gini R, Hyeraci G, Roberto G, Spini A, Lucenteforte E, Agabiti N, Davoli M, Di Domenicantonio R, Cappai G. A Systematic Review of Case-Identification Algorithms for 18 Conditions Based on Italian Healthcare Administrative Databases: A Study Protocol. Epidemiol Prev. 2019 Jul-Aug;43(4 Suppl 2):8-16.