British Isles Network of Congenital Anomaly Registers

BINOCAR Standard Operating Procedure for Data validation and Data sources

Instructions for the Registration and Surveillance of Congenital Anomalies in England and Wales

Authors:
Cath Rounding (CAROBB)
Sylvia Stoianova (SWCAR)

Renewal date: January 2016

Version 1.1 - January 2015

Version number	Date	Comment
1.1	30 January 2015	First version put together by CR, SS and the BINOCAR Management Committee

Contents

Introduction

BINOCAR registers collect variables defined by and transmitted to EUROCAT, additional variables defined by and only transmitted to BINOCAR and local variables, which stay within the local register.

To allow effective analysis of the information collected, each variable needs to be recorded in a clear and consistent way, both within the register and between registers, allowing information to be pooled meaningfully. EUROCAT produces an updated manual with explanations and coding instructions for each variable (Guide 1.4) and BINOCAR extends this work by producing a similar document for the additional BINOCAR variables (BINOCAR extra variables SOP).

The purpose of this document is to bring together documents for validation of data before submission to EUROCAT and BINOCAR and to define how any local data items, which are not covered in other BINOCAR or EUROCAT documents are dealt with.

EUROCAT variables

EUROCAT specifies variables and their coding in EUROCAT Guide 1.4 section 2.2.1, Summary of Variables:

http://www.eurocat-network.eu/content/EUROCAT-Guide-1.4-Section-2.2.1a.pdf

EUROCAT-Guide-1.4
-Section-2.2.1a.pdf

Guidelines for coding anomalies and syndromes are also included in the BINOCAR Standard Operating Procedures for Coding, Classification, Inclusion & Exclusion, Version 4, December 2014.

Coding,
Classification, Inclusic

Guidelines for data validations follow the EUROCAT data validation routines (see EUROCAT Guide 1.4 section 2.5 Data Validation Routines):

http://www.eurocat-network.eu/aboutus/datacollection/guidelinesforregistration/guide1_4

EUROCAT-Guide-1.4
-Section-2.5-1.pdf

EUROCAT data is validated via the EUROCAT Data Management Program (EDMP). As part of the submission process to EUROCAT, data should be imported into EDMP, where an 'Error Check' function runs the specified validation routines and produces an error report. The mechanisms of the EDMP data validation function is explained in the EDMP user guide, Version (6.06), 07/10/2013:

http://www.eurocat-network.eu/content/EUROCAT-Guide-1.4-Section-2.4.pdf

EUROCAT-Guide-1.4 -Section-2.4.pdf

EDMP allows for the validation of multiple cases in one go. Cases may then be corrected locally or marked as valid before exporting the data.

BINOCAR variables

The BINOCAR variables are in addition to those submitted to EUROCAT. These are only sent to the BINOCAR hub and include postcode and demographic information such as ethnic group and smoking status. The BINOCAR extra variables SOP contains basic validation rules for each field which should be adhered to.

Extra BINOCAR variables SOP.docx

Local variables

Registers store variables locally which they do not transmit and are formulated entirely for local use. They may include: place of birth, prenatal diagnostic techniques, previous pregnancies, maternal smoking, maternal alcohol use, age of father at delivery, sources of information, when first reported, aetiology.

NHS number

NHS number, although not transmitted, is essential to the organisation and running of the register, providing the key identifiers.

NHS Number is the only national unique patient identifier allocated to patients having access to NHS services in England and Wales. Using the NHS Number helps share patient information safely, efficiently and accurately linking records across organisations enabling information to be safely transferred across organisational boundaries. Babies are given their own NHS Number at birth to link their healthcare records for life, although there are exceptions.

The Health & Social Care Information Centre (HSCIC) produces standards for NHS numbers, both general standards and specifically for babies:

http://systems.hscic.gov.uk/nhsnumber/staff/guidance

http://systems.hscic.gov.uk/demographics/births

There is no separate record for the fetus. DH guidance is as follows:

From 24 weeks pregnancy all fetuses, whether or not they survive, require an NHS Number.
A baby born showing any signs of life after birth – regardless of gestation period – is a Registerable Birth and will have an NHS Number allocated.

Therefore:

- Guidance and systems do not cater for allocation of an NHS Number to a fetus in the event of, for example, obtaining samples from that fetus
- A stillborn, under 24 weeks - will not be allocated an NHS Number
- A stillborn over 24 weeks - will be allocated an NHS Number
- Fetuses over 24 weeks - will get an NHS Number at birth.

Guidelines and practices vary in the case of conjoined twins. EUROCAT guidelines state that conjoined twins are registered as a single case. Registers however have examples of both babies receiving individual NHS numbers and names as they survived beyond 24 weeks gestation. Conjoined twins data will be submitted as 2 cases to EUROCAT.

Where available (or applicable), all mothers and babies should have their NHS number recorded in dedicated fields on congenital anomaly records to identify duplication and to allow linkage with other data sources.

Registers should request NHS number as a key field on all notification documents and where it is not available, the NHS number should be actively sought.

Finding an NHS Number for a person is achieved by searching for their record on the Personal Demographics Service (PDS) either using the Summary Care Record (for individual cases) or Demographic Batch Service (for bulk files). This process is referred to as tracing and requires an N3 connection. Methods are explained at: http://systems.hscic.gov.uk/nhsnumber

Exceptions and complex cases include:

- Those not registered with the NHS: e.g. some private patients, those resident outside England and Wales
- Armed forces (but this is changing so that armed forces personnel will use NHS numbers)
- Adoptions – previously NHS number was reassigned without a link to the old number but this is changing so that records will be linked and flagged 'sensitive' if appropriate.
- Gender reassignment - There is no link between the patient's original NHS Number and their new NHS Number.

Characteristics of NHS numbers

The NHS number comprises 10 numeric digits. The first nine digits are the identifier and the tenth is a check digit, used to confirm the number's validity.

A verified NHS number is one where the patient's identity has been cross-checked using demographic details on the Personal Demographics Service.

A valid NHS number is one that has the correct format and passes the Number check digit calculation – the Modulus 11 algorithm. With this check there is a minimum of a 90% chance of catching an error from mis-keying (this increases to 100% for most common transcription and transposition errors).

- All NHS numbers should have validation routines to ensure they are exactly 10 numeric digits
- Where possible, NHS numbers should be checked on entry to the database for validity using the Modulus 11 algorithm.
- Ideally, NHS numbers should be verified via the PDS Demographic Batch Tracing Service http://systems.hscic.gov.uk/nhsnumber

Validation of geographical information

All Registers record address and postcode at delivery/outcome, in order to determine area of residence at delivery. Some Registers may also include other addresses and postcodes e.g. place of residence at pregnancy booking, during first trimester, or current residence for child.

All postcodes should be validated to confirm they are valid UK postcodes, and that they match the address. Some Registers' software validates postcodes automatically; otherwise additional geographical software can be used. Alternatively, for small volumes, the Royal Mail has a search facility: http://www.royalmail.com/find-a-postcode.

Anomaly prevalence by geography is calculated using the Office of National Statistics (ONS) birth data based on postcodes supplied by parents at birth registration. Registers' postcodes at delivery (for registerable births) should be checked against matched ONS birth records and amended accordingly to ensure mapping of numerator data (cases) matches denominator data. Registers data notified by NHS services may not contain the most up-to-date address details e.g. hospital casenote labels may not be re-printed following address changes. Birth registration is a statutory process and parents supply address information to the Registrar.

Data validation at the point of data entry

Registers' database systems allow for validation of information at the point of inputting. Error messages are incorporated in the registration database to flag up errors: in the typing of NHS numbers (see NHS number section); when date end of pregnancy recorded before the date start of pregnancy; when a new case entered with a date in the future etc.

Data inputting errors are also minimised by an automatic calculations of some of the data fields derived from entered dates. For instance, father's age at delivery will be automatically calculated once the date of delivery is entered on the database. Other such fields may be:

- Maternal BMI – derived from recorded height and weight
- Gravida – derived from number of previous pregnancies
- Year pregnancy ended – derived from date pregnancy ended
- Gestation length – calculated from date end of pregnancy and LMP
- Age of mother at delivery – calculated from date end of pregnancy and mother's DOB

Diagnostic accuracy of data sources

Congenital anomaly registers receive information from multiple sources such as antenatal clinics, fetal medicine units, maternity units, neonatal units, paediatric departments, physiotherapy departments, clinical genetics departments, cytogenetics, and hospital inpatient data from IT departments. In some cases, the diagnosis is clear and no further follow up is required, or in other cases enough pieces of information are received from different sources to complete the record. Many cases, however, will require additional follow up information. For missing demographic information[1], this is usually quite straightforward but for diagnostic information the situation is

[1] Matching mothers' to babies' records has been challenging. However, an upgrade to the Summary Care Records (SCR) in 2012 had linked mother to baby records.

often less clear. In order for register data to be accurate and comparable, clear definitions of when an anomaly is present are needed.

Registers need consistency in how they make decisions about the validity of information they receive about anomalies, whether the source they receive the information from is considered accurate or whether further follow up is required to confirm the diagnosis. For instance, some routine data sources (e.g. HES) are likely to be inaccurate (or lack specificity) in their diagnostic coding. For instance, electronic downloads from HES data often seem to give Q63.8 instead of (preferred) Q62.0 when they are coding dilated renal pelvis (as anomaly of renal pelvis), rather than equating it with hydronephrosis. Registers may confirm anomaly detail using radiology information systems.

The reliability of the information sources vary depending on the anomaly e.g. a notification received from delivery suite is likely to be reliable for a missing finger, but not for a cardiac anomaly, where assumptions have been made from the antenatal finding. For instance, duct dependent cardiac anomalies, unsuspected prenatally, are unlikely to become apparent until the duct closes approximately 3 days after delivery. An antenatal diagnostic ultrasound scan cannot always diagnose (identify) anomalies such as multicystic dysplastic kidney (MCDK). Occasionally these anomalies are not picked up until the child is older.

To reflect this variation in the diagnostic accuracy of the data sources reported anomalies are given a confirmation status depending on how much detail has been provided and on the reliability of the source. Registers classify each anomaly as confirmed, probable or suspected.

A diagnosis is **confirmed** when there are details of a diagnostic test confirming the anomaly or the data source is considered to be sufficiently reliable. Some anomalies are confirmed antenatally when reported by fetal medicine specialists, such as anencephaly and renal agenesis.

Probable are anomalies considered to be likely but do not have full diagnosis details. Most of these anomalies have been reported by IT departments. Registers try to obtain follow up details to confirm anomalies reported in this way. For example, cleft lip noticed antenatally, at 12 weeks scan for instance, without follow up available is recorded as probable.

Anomalies remain **suspected** when they are reported antenatally, but there was no follow up received by the register at birth. Some conditions reported by IT departments notified with insufficient details will remain suspected until further confirmation is obtained. For example, congenital hydronephrosis (Q62.0) remains suspected if there is only antenatal information available. Q60.2 case will only get confirmed if there are more than one positive postnatal scans.

Only confirmed anomalies are sent to EUROCAT and BINOCAR.

Diagnosis definition

EUROCAT publish an updated guide to definitions of anomalies and also coding tips for anomalies which are difficult to code:
http://www.eurocat-network.eu/aboutus/datacollection/guidelinesforregistration/malformationcodingguides

There are still some grey areas where the situation is unclear and there is scope for inconsistency between registers.

Examples of situations where there may be differing inclusion criteria for anomalies:

- Talipes – reported at birth or only after orthopaedic assessment?
- Hypospadias – reported at birth or only after urology / plastics diagnosis?
- Spina bifida + talipes – EUROCAT only reports on spina bifida. However, there is an argument at local registers' level for recording both anomalies separately
- ASD – What if incidental finding and closes soon after?
- Hydronephrosis – Antenatal hydronephrosis always suspected within the following measurement criteria: a) in isolation: >10 mm (any gestation) and >8 mm (3rd trimester); b) in association with another UM or anomaly: >5 mm; c) also retrospectively code >5 mm if advised of postnatal hydronephrosis. Postnatal measurement criteria: >5 mm and the status is determined by combined results of postnatal scans. Some registers collect information for hydronephrosis with a pelvic dilation < 10 mm but > 5 mm. Cases that do not meet measurement criteria either antenatally or postnatally should be excluded. Register should keep cases that meet antenatal criteria but resolve over ante/postnatal course.

Registers may receive notifications that contradict each other. For example, information received from notifiers antenatally is often unclear or contradictory for cases such as spina bifida +/- Hydrocephalus; cleft lip +/- palate; whether the diagnosis is exomphalos or gastroschisis. In such instances, registers review all the evidence to make an informed judgement, but report anomaly as "probable"; this will mean the case gets sent to EUROCAT. The alternative is to report all possibilities as "suspected", but the case will not go to EUROCAT.

Due to the complexity of the diagnosis definitions a case specific approach should be adhered to. The table below contains examples of some of the most complex cases and what steps should be followed to define the diagnosis.

ICD-10 Code	ICD-10 Condition	Diagnosis Definition
Q79.2	Exomphalos Omphalocele Excludes: umbilical hernia (K42.-)	For antenatal notifications: all cases Suspected below 12/40 gestation, Probable after 12/40 gestation, until confirmation of exomphalos (vs. gastroschisis, cord cyst, umbilical hernia or urachal cyst) by checking on Theatre Records - if confirmed use postnatal confirmation date (e.g. first day check, surgery), or if had been clearly diagnosed by Tertiary Unit antenatally retrospectively confirm by their antenatal ultrasound date.
Q79.3	Gastroschisis (see Q79 for exclusions)	For antenatal notifications: all cases entered as Probable until confirmation of gastroschisis (vs. exomphalos) by checking on Theatre Records - if confirmed as gastroschisis, retrospectively confirm by earliest antenatal ultrasound that clearly diagnosed gastroschisis.
Q05.81	Sacral spina bifida without hydrocephalus - open, aperta, not covered with skin or membrane (see Q05 for inclusions/exclusions)	Probable any antenatal reports, can confirm if associated with banana cerebellum or lemon-shaped head. Can use abnormal AFP result date as date first suspected. 96% of Arnold-Chiari cases include spina bifida, code BOTH (BINOCAR Apr 2005). If severity of spina bifida given this should be entered. Specify whether hydrocephalus ruled out or if status of ventricles not specified. Standardisation of text in form: "Sacral spina bifida, moderate, without hydrocephalus - OPEN" or "Sacral spina bifida, severe, NOS - OPEN". Spina bifida is general term - if type of lesion (e.g. myelomeningocele) known, use this instead. Note standard spelling from coding literature is

		"-cele" not "-coele". See Q06.8 for lipomyelocele and lipomyelomeningocele.
Q05.82	Sacral spina bifida without hydrocephalus - closed, cystica, covered with skin or membrane (see Q05 for inclusions/exclusions)	Probable any antenatal reports, can confirm if associated with banana cerebellum or lemon-shaped head. Can use abnormal AFP result date as date first suspected. 96% of Arnold-Chiari cases include spina bifida, code BOTH (BINOCAR Apr 2005). If severity of spina bifida given this should be entered. Specify whether hydrocephalus ruled out or if status of ventricles not specified. Standardisation of text in form: "Sacral spina bifida, moderate, without hydrocephalus - CLOSED" or "Sacral spina bifida, severe, NOS - CLOSED". Spina bifida is general term - if type of lesion (e.g. myelomeningocele) known, use this instead. Note standard spelling from coding literature is "-cele" not "-coele". See Q06.8 for lipomyelocele and lipomyelomeningocele.

Further points to consider:

- What to do where dates are contradictory?
- Deaths of fetuses before delivery
- The date of the end of the pregnancy is the date of delivery or expulsion
- Date of death as the day IUD was discovered or TOP carried out
- Error in EDMP but at least it highlights where fetuses died prior to delivery.

Sources of missing variables

A case is only recorded if registers have DOB, NHS number and the names of the mother and baby. If any of these demographics are missing, registers search maternity systems and SCR to complete the information. Matching mothers to babies is particularly difficult aspect of this and has been quite common historic challenge. However, upgrade to the national SCR system in 2012 made possible the linkage of mothers and babies records.

When pregnancy details are missing registers consult the maternity systems and if the information there is missing or unsufficient, registers contact the clinician involved with the care of the mother.

Await national data set then list possible sources for each data item.