Data Quality Indicators (updated 08-05-2008)

Introduction

Complex factors influence the quality of data collected by a congenital anomaly registry. How, where and when are diagnoses of congenital anomaly made to residents of the region? How does the registry obtain and verify this information and how does it code it? Registry data are never a "perfect" description of which babies have congenital anomalies in the population and precisely what those anomalies are. Moreover, diagnostic definitions and methods change over time.
Registries may need to make difficult decisions about which types of information to prioritize in order to allocate limited resources, particularly whether to devote resources to collection of exposure information as well as diagnostic information. Some types of information are easier to obtain in some regions than others, depending on specialist referral systems for affected children, confidentiality restrictions, what is recorded in medical notes, availability of computer databases, and willingness to collaborate of key professionals.
EUROCAT's policy is to strive for high quality, accompanied by transparency as to strengths and weaknesses in data quality. Making Registry Descriptions and Data Quality Indicators (DQI) freely available will allow registries to evaluate their performance in relation to other registries, and will allow appropriate interpretations to be made of the results of data analyses. The DQI only have relevance in relation to the objectives of EUROCAT. Different data strengths are needed to participate in timely statistical monitoring in relation to environmental teratogenic exposures, or to evaluate the population sensitivity of prenatal screening for detecting affected pregnancies, or to establish the prevalence of rare genetic syndromes.

The first part of the evaluation of data quality is to read the detailed registry description (available on the EUROCAT website). Key questions to address in evaluating these registry descriptions include the following:
1) Does the registry describe its methods in enough detail? If not, the data emanating from the registry must be regarded as uninterpretable.
2) Is the registry fully population-based? If it covers residents of an area, how does it ensure coverage of residents delivering or seeking paediatric services elsewhere? If it is not entirely population-based, how does it avoid inclusion of selective referrals of high risk pregnancies into registry hospitals?
3) How does the registry ensure coverage of liveborn cases diagnosed after the early neonatal period?
4) How does the registry ensure coverage of late fetal deaths and stillbirths?

5) By what process are terminations of pregnancy following prenatal diagnosis identified in the population?
6) What specialist services are accessed for information, and are these services used for case identification or only for follow-up of known cases?

The second part of the evaluation of data quality is to look at the data generated by the registers.
The following sections provide a set of Data Quality Indicators and a description of completeness of registry data transmitted to the EUROCAT Central Database. The DQI of a particular registry can be compared to the EUROCAT average. Strong deviations on either side of the average should be examined. The set of data quality indicators has been produced under the
headings:
Completeness of case ascertainment
Accuracy of Diagnosis
Completeness of Information on core and non-core EUROCAT variables
Timeliness of data transmission
Availability of Denominator Information

This first ever set of DQI was discussed in Poznan June 2005.

List of Data Quality Indicators (1999-2003)

Ascertainment

• Total congenital malformations prevalence = All cases in malformation chapter ICD10 (Q chapter) or ICD9 (range 740-759), including outside malformation codes (D821, D1810, P350, P351, P371, D215, 27910, 2281, 7710, 7711, 77121). EUROCAT minor anomalies are excluded.

• Ratio of Spina bifida to Anencephalus.

• Prevalence NTD = Neural Tube Defects

• Prevalence selected cardiac malformations
Hypoplastic left heart, Transposition of Great Vessels, Tetralogy of Fallot, Coarctation of aorta or Common arterial truncus.

• Prevalence selected postnatal diagnosed malformations
Corpus callosum anomalies, Cataract, Coarctation of aorta, Hirschprung's disease, Unilateral renal agenesis, or Craniosynostosis.

• Prevalence non-chromosomal syndromes
ICD/BPA9 codes: 27910, 755500, 755501. 755551, 75581, 75601, 756041, 756050, 75606, 756110, 75640, 756421, 75643, 756540, 756441, 756442, 756443, 756512,

75652, 756540, 756551, 756581, 75680, 75683, 75685, 756883, 756885, 75712, 757303, 75735, 759340, 759611, 759802, 759804, 759807, 759820, 759821, 759822, 759823, 759826, 758927, 759828, 759840, 759841, 759842, 759845, 759846, 759860, 759870, 759872, 759874, 759875, 759890, 759892, 759895, 759896, 759897, 759898, 759899, 7710, 7711, 77121

ICD10 codes: D821, P350, P351, P371, Q4471, Q6190, Q7402, Q7484, Q751, Q754, Q7581, Q761, Q771, Q772, Q775, Q776, Q7780, Q7800, Q781, Q782, Q783, Q7840, Q796, Q7982, Q823, Q8580, Q8581, Q8582, Q8583, Q860, Q861, Q862, Q8680, Q8682, Q87, Q8934, Q936, Q992,
Excluding: Q8703, Q8706, Q8708, Q8724

• Prevalence malformed fetal deaths calculated using total births.

Accuracy of Diagnosis

• % of possible multiple malformations in database excluding chromosomal or syndrome cases
using the EUROCAT flow-chart for multiple malformations

• % fetal deaths and terminations of pregnancy with post-mortem examination performed

• % of chromosomal cases with a karyotype performed

• % of non-chromosomal / non-syndrome multiple malformation cases in database with known karyotype
Using the EUROCAT flow-chart for multiple malformations

• Prevalence of selected Q-BPA extension codes (restricted to registries that use ICD10 coding).
Selected Q-BPA codes = Q00.00, Q00.20, Q04.00, Q04.35, Q21.10, Q21.21, Q25.10, Q25.11, Q26.20, Q33.80, Q37.10, Q39.11, Q44.20, Q61.41, Q64.20, Q71.31, Q89.80

• Prevalence of selected Q-chapter unspecified codes (restricted to registries that use ICD10 coding).
Selected unspecified codes = Q04.9, Q05.9, Q24.9, Q33.9, Q43.9, Q54.9, Q63.9, Q74.9, Q79.9, Q89.9, Q99.9

Completeness of information

Completeness of information describes the amount of **complete valid data** transmitted to Central Registry (eg. “Not known” values, invalid values, or missing/blank fields are counted as incomplete information).

• There are 7 core variables (Sex, Nbrbaby, Type, Weight, Gestlength, Survival, Agemo).

• There are 39 non-core variables (Place, Birthord, Nbrmalf, Death_date, Datedisc, Whendisc, Agedisc, Condisc, Amnio, Chorvilsam, Othertech, Karyo, PM, Datemo, Residmo, SA, IA, LB, SB, Totpreg, Occupmo, Assconcept, Illbef, Illdur1, Habit1, Unusexp, Drugs1, Datefa, Agefa, Occupfa, Mckusick, Modetrans, Consang, Prevsib, Sib1, Sibanom, Moanom, Faanom)

Please note that there are some changes to the list of core and non-core variables for cases born pre-2005 and post-2005.

• Written text for syndrome and malformations