3.1. Overview of EUROCAT Approach to Coding and Classification

The purpose of **coding congenital anomalies** in EUROCAT registries is to summarise the non-standardised written text in such a way that the data can be analysed for surveillance and research purposes. Registries can encourage but can rarely impose the use of standard definitions and diagnostic tests and clinical follow-ups, and the coding system must allow for different levels of precision and accuracy of information provided by clinicians.

The purpose of **a "classification" system**, in the sense that we use this term here, is to group anomalies which share aetiologic or clinical characteristics. There is a balance to be struck *a)* between "lumping" together heterogeneous sets of anomalies and "splitting" so finely that there are few cases in each group and *b)* between creating groups based on great precision and accuracy of diagnosis and coding and creating groups which take into account what can be realistically found in medical records and regional or national databases for most cases.

The standard dataset for each case (see Chapter 2.2) allows for the text description and coding of up to eight malformations and one syndrome. Where there are more than eight malformations, additional malformations can be added in the text variable for the 8th malformation. When there is more than one syndrome, the second syndrome can be coded in the first malformation variable.

Malformations are coded to ICD10 with the British Paediatric Association (BPA) one-digit extension. Syndromes are also coded to ICD10-BPA. An OMIM (McKusick) code can be given in addition but should be coded with caution and expertise. A specific EUROCAT guide to the coding of syndromes gives more detail ("EUROCAT SYNDROME GUIDE: Definition and Coding of Syndromes", see link). Where ICD9-BPA codes are given in this Guide, this is solely for the purpose of analysing the EUROCAT database from 1980 to 2004. ICD9 codes should no longer be used.

"All anomalies": The core set of congenital anomalies to be registered by all member registries are structural malformations and chromosomal anomalies diagnosed in the fetus, baby or child. Particular attention is paid to complete ascertainment of those anomalies usually diagnosed in fetal life or the first year of life. In order to set clear boundaries to this group and achieve comparability of prevalence between registries, the count of cases with all anomalies includes all cases with one or more codes in the Q chapter of ICD10 and a very limited set of conditions coded outside the Q chapter, as specified in the definition of subgroups given in Chapter 3.3 (see subgroup "all anomalies").

In addition, registries may register other congenital conditions, including congenital neoplasms (ICD10 C and D codes), and congenital endocrine (E), metabolic (E), immunologic (D), and haematologic (D) conditions. These are not included in the prevalence of "all anomalies". Registries should make clear in their registry description which non-Q conditions they register.

Impairments of function with a partially or wholly prenatal origin, such as cerebral palsy, autism or developmental delay/intellectual disability, should not be transmitted to EUROCAT unless in association with specified structural malformations.

Cases with only minor anomalies as specified on the EUROCAT list of minor and unspecified anomalies for exclusion (Chapter 3.2) should not be transmitted to EUROCAT. Minor anomalies should be described in the text, coded and transmitted to EUROCAT when they are in association with major anomalies. Cases with only minor codes will be automatically excluded by the computer. Some minor anomalies as given in Chapter 3.2 do not, however, have specific codes and cases with such isolated anomalies must always be recognised and excluded at a local level on the basis of the text description.

The EUROCAT subgroups as defined by their ICD10 codes (Chapter 3.3) are subgroups for which prevalence information is routinely produced. These subgroups are also the subject of routine statistical monitoring for trends and clusters in time. Subgroups have been defined according to one or more of the following criteria:

a) Larger heterogeneous subgroups which show the relative health burden of anomalies in different organ systems,

b) Subgroups which balance aetiologic homogeneity with the level of diagnostic specificity which can reasonably be expected by European registries,

c) Subgroups which are relevant to health service provision, including prenatal diagnosis,

d) Subgroups which are well defined and clinically diagnosed with a good level of consistency across Europe, and where specific codes are available,

e) Subgroups that are consistent with the hierarchical classification of ICD10,

f) Subgroups of reasonable frequency such that a yearly European prevalence can be meaningful.

g) Only major anomalies (i.e. not on the list for exclusion in Chapter 3.2) are allocated to subgroups, however where the same code specifies both a minor and major anomaly, minor anomalies may be included in subgroups.

Prevalence is published on the website both including and excluding cases with genetic disorders. Cases with genetic disorders are always excluded from the statistical monitoring of structural anomalies (analysis of clusters and trends).

All prevalence estimates and counts for subgroups are based on cases, not malformations. Thus, a baby with a VSD and pulmonary valve stenosis will be counted ONCE in “all anomalies”, ONCE in “congenital heart defects”, ONCE in “VSD”, ONCE in “pulmonary valve stenosis”. A baby with encephalocele and multicystic renal dysplasia will be counted once in the count of “all anomalies”, once in the count of cases with central nervous system anomalies, once in the count of cases with neural tube defects, once in the count of cases with encephalocele, once in the count of cases with multicystic renal dysplasia and once in the count of cases with congenital anomalies of kidney and urinary tract. It follows that the number of cases in different subgroups CANNOT be added together to find the total number of cases, as **one case can be counted in more than one subgroup**. A higher prevalence of subgroups can be expected in areas where more detailed coding of multiply malformed babies is undertaken.

Multiply malformed cases are the subject of separate statistical monitoring with the purpose of identifying teratogenic exposures that cause patterns of multiple malformations. Cases which are likely to be multiply malformed are identified using the hierarchical computer multiple congenital anomaly algorithm given in Chapter 3.4. A manual review of the identified potential multiply malformed cases (approx. 10%) will be done by the geneticists of the EUROCAT Coding and Classification Committee before statistical surveillance.