# MEDICAL GENETICS IN PRACTICE

# An aetiological classification of birth defects for epidemiological research

## D Wellesley, P Boyd, H Dolk, S Pattenden

.......................................................................................................................

See end of article for authors' affiliations
......................

Correspondence to:
Dr D Wellesley, Head of Prenatal Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital, Coxford Road Southampton SO16 5YA; dgwsoton.ac.uk

**Background:** Congenital anomaly registers collect data on antenatally and postnatally detected anomalies for surveillance, research, and public health purposes. Each anomaly is coded using the International Statistical Classification of Diseases and Related Health Problems (ICD-9/ICD-10) based on body systems, allowing accurate comparisons between registers for individual anomalies. When commencing an environmental, epidemiological study, it became clear to us that there is no standard classification that takes aetiology into account. This paper describes a new classification for use in studies addressing aetiology.
**Method:** A classification system was evolved and piloted using cases in a study of geographical variation in congenital anomaly prevalence.[1] Cases that were difficult to categorise were noted, and after discussion with a team of experts, the classification was adjusted accordingly.
**Results and conclusion:** A robust, hierarchical method of classifying birth defects into eight categories has been produced, for use at source of data registration in conjunction with, but independent of, ICD coding.

S ignificant progress has been made into understanding the aetiology of birth defects over the past 20 years. The detection of chromosome microdeletions has explained the aetiology of some non-Mendelian syndromes and the identification of single gene mutations has shown or confirmed the Mendelian basis for others. However, in spite of advances in new genomic technology, many birth defects have unknown aetiology and there is concern about environmental causes.

Epidemiological studies looking into the causes of birth defects have not always taken different aetiologies into account, clumping together all cases of a specific defect for analysis. For example, diaphragmatic hernia could be an isolated anomaly or associated with a chromosome error such as trisomy 18 or might be part of the autosomal recessive Fryns' syndrome. To derive a valid conclusion from such a study it is necessary to ensure that aetiology has been taken into account.

Population based birth defect registries are a valuable resource for studies of environmental and genetic aetiology. ICD-10[2] has become established as the international coding system of choice for individual specific anomalies,[3] and is used by most registers for this purpose.

Classification of cases into anomaly subgroups is more difficult and is dependent upon the intended use of the data. Surgeons have classically used a morphological approach, whilst geneticists prefer a system based on inheritance, as in McKusick's *Mendelian Inheritance in Man*,[4] which provides a number for every described single gene disorder, using an inheritance based tree structure. For dysmorphologists, Jones[5] proposed a system reclassifying all birth defects into malformations, deformations, and disruptions, incorporating aetiological links that are sometimes speculative. None of these, however, fulfils the requirements for epidemiological research requiring grouping by aetiology.

For epidemiological studies it is necessary to classify defects according to presumed aetiological commonality, allowing further investigation of those cases where an, as yet unknown, environmental element could be important.

In 2000, commencing an investigation into the patterns of geographic variation in congenital anomaly rates across five geographical areas of Britain, it became clear that a robust system of classification was necessary.

The aim of this paper is to describe a system for the classification of birth defects by aetiology for epidemiological research, to be used by congenital anomaly registers at the time of case registration. Classification would thus take into account information available at the time of registration, which may include family history, and diagnostic and laboratory test data in addition to the anomalies.

The basic requirements for such a system include: (*a*) ease of use by people of different scientific backgrounds; (*b*) all anomalies covered—that is, no miscellaneous group; and (*c*) a method of categorising those with no specific diagnosis.

## METHODS

The study population comprised 845 000 births to mothers resident in five geographically defined areas of the UK that have congenital anomaly registers, covering the period 1991–1999. Well defined, significant anomalies were included. In all, 10 844 cases were analysed.

A hierarchical system of classification was drafted in which cases could be classified to one category only, the highest in the list of categories applicable. This was piloted (by DW and PB) to categorise the 10 844 cases in the geographical heterogeneity study.[1] Thus, it was possible to highlight areas for adjustment, draw attention to the cases that were difficult to classify, and test the overall robustness of the classification system. Some modifications were then applied. For further validation of the groups, the proposed system was presented at the 2002 British Isles Network of Congenital Anomaly Registers (BINOCAR) meeting for those actively involved with running or managing congenital anomaly registers, and at the Tenth Manchester Birth Defects meeting, attended by clinical geneticists, for further consideration, with particular emphasis on the aetiology of specific defects. In addition, a list of ''difficult to categorise'' cases was offered to a panel of experienced geneticists for their opinion. A consensus was

.......................................................................

reached, and these results have been incorporated into the guidance lists, which have been provided for many categories to improve consistency.

## RESULTS

Table 1 gives the final classification categories. The flow chart illustrates the route through the hierarchical system. Table 2 lists the diagnoses to be included in the microdeletion (MD) category; table 3 the list of teratogens (T), and table 4 the new dominant (ND) mutations.

## Analysis

In order to use this classification system for epidemiological studies of potential environmental influence we recommend: (*a*) those cases in the teratogen and familial categories are excluded as of known aetiology; (*b*) the chromosome, microdeletion, new dominant mutation, and syndrome subgroups are mutually exclusive categories, separately analysed without reference to individual component anomalies. an example here would be a heart defect in Down's syndrome which would only be analysed in the chromosome group; and (*c*) malformations in the multiple category are counted in each of the different anomaly subgroups represented, as well as in a combined ''M'' category.

Although designed with environmental epidemiology in mind, these categories may also be used for wider ranging studies such as new dominant mutation rates, where consistent categorisation will allow accurate comparisons between registers and regions.

---

**Table 1** Categories for the classification of birth defects in hierarchical order

| | |
|---|---|
| C: Chromosome | For microscopically visible, unbalanced chromosome abnormalities. |
| MD*: Microdeletion | For all submicroscopic chromosome abnormalities including microdeletions, uniparental disomy and imprinting mutations. |
| T*: Teratogen | For known teratogens and prenatal infections. |
| ND*: New dominant | For new dominant mutations. |
| F: Familial | For familial disorders not included as a new dominant. |
| S: Syndrome | For recognised non-familial, non-chromosomal syndromes. |
| I: Isolated | For isolated anomalies. |
| M: Multiple | For multiple anomalies and associations. |

*List provided. C: includes trisomies (+ mosaics), triploidy, visible deletions such as 4p (Wolf-Hirschorn), duplications, insertions, and unbalanced translocations. Microdeletions in category ''MD'' are excluded. MD: all submicroscopic chromosome abnormalities. These disorders may be reported to a register in different ways depending on the source of the data, thus any chromosome error or syndrome name mentioned in table 2 should be coded MD. T: cases with a good history of teratogen ingestion, or a specified infection, as well as anomalies strongly consistent with the exposure (table 3). ND: autosomal dominant conditions where the new mutation rate is ≥90% and the diagnosis easy to make at birth, as detailed in table 4. If there is known to be a family history of the condition it should be coded as ''F'' but ''ND'' is the default category for these conditions. F: all non-chromosomal, familial conditions not included in ''ND'', such as tuberous sclerosis, Meckel-Gruber syndrome, autosomal recessive polycystic kidney disease, fragile X syndrome. S: all non-chromosomal, non-Mendelian syndromes not included in one of the previous categories, such as Kabuki syndrome. I: all isolated anomalies not included in one of the above categories. Included here are (*a*) single anomalies such as gastroschisis, talipes, or cleft lip; (*b*) more than one anomaly from the same system such as a VSD and coarctation, or polydactyly of hands and feet; and (*c*) more than one anomaly as part of a sequence such as spina bifida with talipes or hydrocephalus, lung hypoplasia and renal agenesis. M: all that remain should be multiple, unrelated anomalies from more than one system with no unifying diagnosis and recognised associations that are not regarded as a syndrome (for example, VATER, CHARGE, MURCS).

---

**Table 2** Diagnoses to be coded MD

| Chromosomal/genetic error | Syndrome names |
|---|---|
| 22q11 deletion | Di George, velocardiofacial or Schprintzen syndrome |
| 15q11 deletion, maternal or paternal disomy, imprinting mutations | Prader-Willi or Angelman syndrome |
| 7q11 deletion, elastin mutation | Williams syndrome |
| 11p13 deletion, PAX 6 or WT1 mutation | WAGR, Aniridia-Wilms' tumour |
| 11p15 duplication, paternal isodisomy, imprinting mutations | Beckwith-Wiedemann syndrome |
| 20p12 deletion or JAG1 mutation | Alagille syndrome |
| 17p11 deletion | Smith-Magenis syndrome |
| 5q35 deletion, NSD1 mutation | Sotos syndrome |

---

**Table 3** Teratogens and prenatal infections included in ''T'' and where observed defects are strongly associated with the teratogen

Drugs
    Abortifacients
    ACE inhibitors
    Alcohol
    Anti-epileptics
    Cocaine and other illicit drugs
    Cytotoxics
    Diabetes, poorly controlled
    Folic acid inhibitors
    Lithium
    Severe maternal illness
    Thalidomide
    Vitamin A analogues
    Warfarin
Infections
    Cytomegalovirus
    Herpes simplex
    Parvovirus
    Rubella
    Toxoplasmosis
    Varicella

---

**Table 4** New dominant mutations included in ''ND''

Achondroplasia
Apert's syndrome
Thanatophoric dysplasia
Campomelic dysplasia
Osteogenesis imperfecta types II* or III

*Type IIc is recessive and ideally should not be included here but if the type of OI II is not known, it should be included in this category.

## DISCUSSION

In order to make progress in identifying new aetiological factors, epidemiological research needs to take into account new knowledge in clinical genetics. Without careful attention to classification based on knowledge of the causes of birth defects, epidemiological studies may fail to identify environmental teratogens because of the inclusion of cases with the same defect but different aetiologies. We have tested this classification system in a study of geographical and sociodemographical variation in the rates of congenital anomalies.[1]

Although it is possible to categorise defects appropriately for a specific research study, it is more accurate and less time consuming to do so at source when entering and coding each anomaly, using locally available information such as family

history and test results. A classification system needs to be simple to use by all those involved in registering anomalies, to be consistent and reproducible. It also needs to be appropriate for the circumstances. There is no doubt that a more detailed system using case specific genetic information would be more accurate, but this would not be operable by local congenital anomaly registers. If future research is to be immediately comparable, the same system must be usable by all.

## Microdeletions

The categorisation of syndromes now known to be due to a chromosome microdeletion, uniparental disomy or an imprinting mutation was widely discussed during the pilot phase. The main priority was to keep all cases with the same diagnosis in the same category. The option of categorisation under ''syndrome'' was considered, but some concerns were raised. If the source of a case is the cytogenetic laboratory, the diagnosis will be given as, for instance, 22q11 deletion and hence coded C. A similar case from the cardiologist might be reported as Di George or velocardiofacial syndrome and coded S. Conditions such as Prader-Willi syndrome may be due to a paternal microdeletion, maternal isodisomy, or an imprinting mutation. Some will be reported as 15q11 deletion by the laboratory and others as Prader-Willi syndrome, with no mention of the underlying mechanism. Other diagnoses such as Alagille syndrome may have a chromosome deletion, a positive family history or neither and thus could be coded as C or F or S.

Recent research has suggested that imprinting mutations may be more common following assisted reproduction,[12–14] involving mechanisms of which we currently have little knowledge. Further similar research questions are inevitable. It is therefore important to be able to retrieve such cases with ease. For these reasons, a separate category, MD, has been created for all of these cases. A disadvantage of this category from a purely genetic perspective is that it pools basic genetic mechanisms (imprinting, uniparental disomy, and chromosome microdeletions) that probably do not have the same underlying aetiology. A compromise had to be reached so that cases with the same diagnosis, but reported in different ways (for example, di George syndrome as: 22q11 deletion, or di George, Schprintzen, or velocardiofacial syndromes) were not lost, and to provide for the fact that, in many cases, the basic genetic error will not be known by the register. The MD category keeps all these cases together for potential future research and offers clear guidance as to their coding. Finer subdivisions can later be made depending on the specific research question. A list of conditions to be included in this category is provided (table 2), to be updated annually, taking into account new knowledge that may move a condition from one category to another, for example, from S to MD. In theory, these cases could also be classified under C but it would not be possible, or desirable, to produce an inclusion list for all chromosomal errors. As a list seemed the only way to ensure inclusion of all like cases, a specific category was provided.

## Teratogens

This is often a difficult category for diagnosticians. For the purposes of this classification system, in situations where, for environmental epidemiological studies, cases are excluded from further analysis if aetiology is known, only those cases where the diagnosis is clear should be categorised as T. Where there is uncertainty, M or I would be preferable.

## New dominants

One of the most controversial categories is ND, for new dominant mutations. These cases are small in number but of importance, and have been promoted for decades as the monitor of mutation rates in populations.[6–8]

For this category, consistency was deemed of the greatest importance. Some registers have access to good family history data and this would allow accurate classification of cases such as tuberous scerosis or Noonan's syndrome (between F and ND); other registers have poorer access. In the absence of such data, allocation is likely to vary between registers. It was therefore decided to use this category in a limited way accepting only conditions that are usually (≥90%) new mutations, reliably diagnosable at birth. A known family history would supersede these assumptions, but in the absence of this information, a new mutation would be assumed. This category will allow an easy comparison of the new mutation rate between registers.

Recent research has shown that new mutations in Apert's syndrome, achondroplasia, and (probably) thanatophoric dysplasia are exclusively paternal, and related to paternal age.[9–11] This does not invalidate the value of new dominant mutations as a measure of pro-mutagenic processes, but suggests it would primarily assess paternal new mutations and that paternal age should be taken into account when making population comparisons.

## Syndrome

One of the more common problems identified, by discussion, at the BINOCAR conference is loose usage of the word ''syndrome''. This term brings together many conditions with very diverse aetiologies, such as Down's syndrome, Di George syndrome, Smith-Lemli-Opitz syndrome, and Kabuki syndrome. The value of the hierarchical system is that it allows each of these to be categorised correctly before the offer of category S.

The lists provided for specific categories err on the side of caution. Assumptions have been made that some registers will have difficulty obtaining information about family history (relevant to new dominant cases in particular), association diagnoses (currently included within the M group) and other issues. If this proves not to be the case, the ND category, for example, could be extended to include conditions such as tuberous sclerosis, and Pfeiffer, Crouzon, or Noonan's syndromes. Several registers within the UK BINOCAR network have agreed to pilot this system at source with a planned annual review. After this first annual appraisal by the authors, the BINOCAR coding group will meet annually to review the categories in the light of new diagnoses or tests. Problem cases can be submitted to this group and changes made as necessary at this time. Thus, consistency will be maintained.

Rassmussen et al[15] recently published a classification system for the US National Birth Defects Prevention Study. It is specific to the cases included in that study for use by medical geneticists when classifying cases, but its basis is very similar to that described here. A hierarchical system with chromosome, genetic, teratogen, syndrome, isolated, and multiple categories is reported. The similarity between these two independently derived systems would seem to add weight to our proposed classification. We also extend the system for use prospectively at the point of data entry in registers.

We accept that this proposed classification is not perfect and has required compromise, but believe it to be significantly better than using ICD codes alone.

As a work in progress, there will be the opportunity for continual improvement.

In summary, an aetiological classification of birth defects has been developed for use by registers at the point of data entry. We believe this to be a simple, robust and reproducible
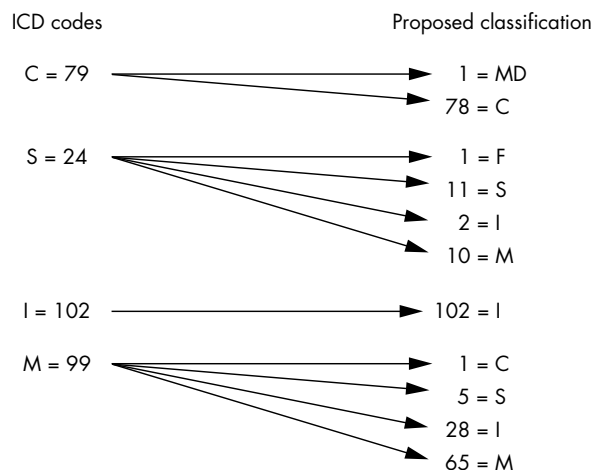
**Figure 1** The categorisation of exomphalos using the proposed classification system compared with ICD codes.
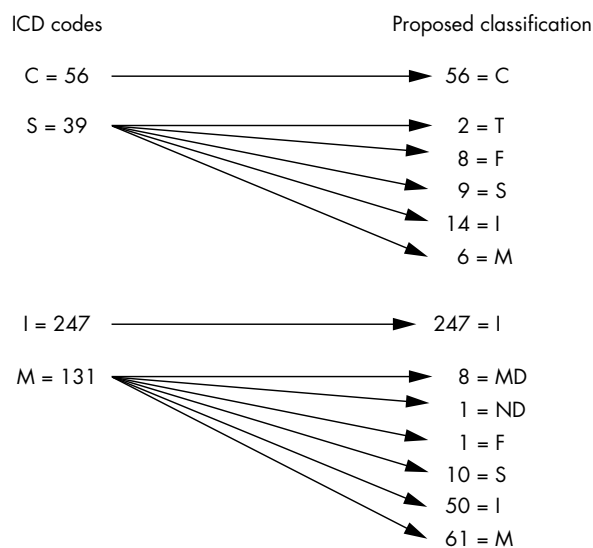


**Figure 2** The categorisation of cleft palate using the proposed classification system compared with ICD-10.

system for use in future epidemiological and aetiological research.

. . . . . . . . . . . . . . . . . . . .

**Authors' affiliations**
**D Wellesley,** Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton SO16 5YA, UK

**P Boyd,** CAROBB National Perinatal Epidemiology Unit, University of Oxford, Old Campus Road, Headington, Oxford OX3 7LG, UK
**H Dolk,** Faculty of Life and Health Sciences, University of Ulster at Jordanstown, Newtownabbey, Co Antrim BT37 0QB, UK
**S Pattenden,** Environmental Epidemiology Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK

Competing interests: none declared

## APPENDIX 1

### PILOT STUDY
There are currently two analytical approaches available to environmental epidemiological researchers. The most basic, used by many studies in the past, groups all cases of a particular anomaly together. Hence the 304 cases of exomphalos would have been analysed as one group instead of: 79 C, 1 ND, 1 F, 16 S, 132 I, and 75 M if the proposed classification was used. Using I as the default category, 57% would be re-classified.

For cleft palate, comparing our classification with a single group analysis would result in a 34% re-classification as follows: 56 C, 8 MD, 2 T, 1 ND, 9 F, 19 S, 311 I, and 6 M.

More recently, some researchers have used ICD codes to separate out cases with a chromosome aetiology and cases with isolated versus multiple anomalies. Figs 1 and 2 compare this method of categorisation with the proposed classification. Comparing these approaches resulted in re-classification in 16% of the exomphalos cases and 21% of those with a cleft palate.

## REFERENCES
1 **Dolk H**, Abramsky L, Armstrong B, Botting B, Boyd PA, Dunn C, Grundy C, Jordan H, Pattenden S, Physick N, Rankin J, Stone D, Vrijheid M, Wellesley D. A study of geographical variation in overall rates of congenital abnormalities and the rates of specific abnormalities: Final report to the Department of Health. DH/DETR/Environmental Agency Joint Research Programme on the possible Health Effects of Landfill Sites. 2003.
2 **WHO**. *International statistical classification of diseases and related health problems,* 10th revision ed, vol. 3, 1994.
3 **Rasmussen, S, C**. Moore. Effective coding in birth defect surveillance. *Teratology* 2001;**64**(suppl 1):S3–7.
4 **McKusick V**. *Online mendelian inheritance in man.* Available at: http://www.ncbi.nlm.nih.gov/Omim/searchomim.html.
5 **Jones K**, M Jones, J Fisher. A developmental approach to the classification of birth defects. *Ann Plast Surg* 1985;**14**:135–41.
6 **Chandley A**. On the parental origin of de novo mutation in man. *J Med Genet* 1991;**28**:217–23.
7 **Czeizel A**. Population surveillance of sentinal anomalies. *Mutat Res* 1989;**212**:3–9.
8 **Risch N**, Reich EW, Wishnick MM, McCarthy JG. Spontaneous mutation and parental age in humans. *Am J Hum Genet* 1987;**41**:218–48.
9 **Moloney DM**, Slaney SF, Oldridge M, Wall SA, Sahlin P, Stenman G, Wilkie AO. Exclusive paternal origin of new mutations in Apert syndrome. *Nature Genet* 1996;**13**:48–53.
10 **Sapienza C**. A paternal wash in Apert syndrome. *Nature Genet* 1996;**13**:9–10.
11 **Goriely A**, McVean GA, Rojmyr M, Ingemarsson B, Wilkie AO. Evidence for selective advantage of pathogenetic FGFR2 mutations in thèmale germ line. *Science* 2003;**301**:643–6.
12 **Cox GF**, Burger J, Lip V, Mau UA, Sperling K, Wu BL, Horsthemke B. Intracytoplasmic sperm injection may increase the risk of imprinting defects. *Am J Hum Genet* 2002;**71**:162–4.
13 **DeBaun M**, Niemitz E, Feinberg A. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *Am J Hum Genet* 2003;**72**:156–60.
14 **Gosden R**, Trasler J, Lucifero D, Faddy M. Rare congenital disorders, imprinted genes, and assisted reproductive technology. *Lancet* 2003;**361**:1975–7.
15 **Rasmussen SA**, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA; National Birth Defects Prevention Study. Guidelines for case classification for the National Birth Defect Prevention Study. *Birth Defects Res* 2003;**67**:193–201.

**JMG**

# An aetiological classification of birth defects for epidemiological research

D Wellesley, P Boyd, H Dolk and S Pattenden

---

Updated information and services can be found at:

http://jmg.bmj.com/content/42/1/54

---

|  | *These include:* |
|---|---|
| **References** | This article cites 12 articles, 2 of which you can access for free at:<br>http://jmg.bmj.com/content/42/1/54#ref-list-1 |
| **Email alerting service** | Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article. |

---

| **Topic Collections** | Articles on similar topics can be found in the following collections<br>Epidemiology (630) |
|---|---|

---

**Notes**

To request permissions go to:

http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:

http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:

http://group.bmj.com/subscribe/