

1

The development of registries to support birth defect research

Birth defects – congenital anomalies (CA) – are a major cause of infant mortality and childhood morbidity, affecting 2-3% of all babies. CA are also responsible for large numbers of embryonic and fetal deaths. The appropriate treatment of liveborn infants with significant birth defects makes heavy demands on health care resources. The epidemiological approach to birth defects has been the backbone of research into their causes. Hypotheses about possible causative agents may arise from many different sources – from the observations of astute physicians, from experimental animal teratology, from epidemiological studies themselves – but epidemiological techniques are usually necessary to test these hypotheses.

1.1 Type of registry

Whenever a system for the registration of birth defects is to be established (WHO, 1998, 2003), a number of decisions has to be made. These will be influenced by the purposes for which the registry is being established, the resources available and the practicability of the scheme in relation to geographical, administrative, cultural and other relevant factors. Among the most important decisions are the following:

- **Will the registry be population- or hospital-based?** A population-based registry records data relating to all births² to mothers resident within a defined area, irrespective of where the birth takes place. A hospital-based registry – based in one or many hospitals – records defects in babies born, irrespective of where the mother lives. Birth prevalence rates derived from hospital-based data are more liable to bias than are population-based rates.

² “Births” include live births and stillbirths and, where appropriate, pregnancies terminated because of prenatally diagnosed birth defects.

- **What sources of ascertainment will be used?** The more sources of information that can be used, the larger the proportion of cases that will be recorded. The basic source of information is usually the birth record, but this is often incomplete, particularly for congenital heart defects. Additional sources of information include hospital admissions, neonatal surgery records, attendance at special clinics (e.g., paediatric, cardiology) or child health clinics, postmortem reports and school health records.
- **Will all defects be recorded or only a selected list?** It is usual to exclude a number of minor anomalies, largely because of the difficulty of distinguishing between minor anomalies and normal variants, and because of the variability in the reporting of such “defects”. With these minor exceptions, it is desirable to record all defects. A monitoring programme that excludes any major defect runs the risk that a new teratogen may cause a defect which is not being recorded.
- **Will stillbirths be included?** If the objectives of the registry include the highest possible level of ascertainment, it is highly desirable that stillbirths be included, preferably with a postmortem report from a paediatric or perinatal pathologist.
- **Will termination of pregnancy following prenatal diagnosis be included?** As an increasing number and variety of congenital anomalies are being dealt with by terminating the pregnancy (where the law allows it and where the facilities exist), it becomes increasingly important that these defects should be included. For example, in many centres the pregnancies are being terminated for more than 90% of the cases of anencephaly and more than 50% of those with spina bifida and Down syndrome. By contrast, it is virtually impossible to collect information routinely on defects in spontaneously aborted fetuses; these are normally, by convention, excluded from registration.
- **For how long after birth will data be collected?** For obvious external defects ascertainment should be virtually complete at birth although, in practice, recognition is not always translated into notification. For some internal defects, notably congenital heart disease and anomalies of the genito-urinary tracts, diagnosis at birth is often seriously incomplete. A one-year follow-up is a common and useful compromise.
- **Will any control group of healthy babies be recorded to allow for comparison of demographic and other factors between the parents of healthy infants and those with birth defects?** The value of this is debatable. When the need arises for a case-control study of

a possible teratogenic risk factor, it is often the case that the variable to be studied has not been included in the predetermined list of factors to be recorded.

For the purposes of each individual scheme, decisions on these and other issues must be made to suit the circumstances and purposes. As long as the local rules and definitions are agreed upon, known and adhered to, useful comparisons can be made between one time period and another. However, if comparisons are to be made between one registry and another, or between one country and another, the greatest care must be taken to ensure that like is being compared with like.

Prevalence rates
vary from month to
month and from
year to year ...

Differences in birth defect prevalence between places and over time are inevitably affected by the time scale and extent of the geographical area examined. If a small area is studied, the number of cases of individual defects will be small. Significant time trends, whether increases or decreases, will be rare. If a large area is studied, case numbers will be higher and significant time trends will be easier to detect. However, it then becomes necessary to determine whether the trend is affecting the whole area or only parts of the area, whereupon the problem of small numbers arises again. However, the study of small areas increases the chance of identifying associations with specific environmental risk factors.

Just as studies of small areas have limitations, it is unlikely that significant time trends will be detectable from observations limited to a short period of time. In both cases, the denominator (the total number of births) is small. The longer the period over which baseline observations are made, the more confidence can be placed in the significance of changes in prevalence. However, comparing later rates with previous rates is not without risk. A careful analysis of previous rates is necessary to be sure that no clusters or trends were already present. It is important to note that ascertainment techniques change over time.

As most birth defects appear to be random events, prevalence rates vary from month to month and from year to year. To establish reasonably stable baseline rates, registries covering relatively small numbers of births (e.g., fewer than 20 000 per year) may need a 10-year baseline; large registries (e.g., more than 100 000 annual births) may need 5 years. In the first year of operation of any new registry, when lines of communication are being established, ascertainment is usually less complete than in subsequent years.

1.2 Birth-defect surveillance and support of research using registries

Prevalence information is a basis for comparison and collaboration ...

As the main purpose of most registries is to provide information on the prevalence of birth defects, collection of good prevalence data is likely to be expensive and difficult. Depending on logistic opportunities, registries range from those with (assumed) complete coverage of all births of a defined population to hospital-based registries that cover populations that are more vaguely defined. When comparing prevalence from such different registries it is important to acknowledge differences – both in the ways birth defects are diagnosed and reported, as well as in the reliability of numerators of prevalence estimates.

The traditional paradigm of prevalence data surveillance has been that detection of a particularly high prevalence in one particular population, or a sudden increase in prevalence over time, could help identify possible environmental causes (Lie et al., 1991). Prevalence information therefore serves as a basis for comparison and collaboration. Examples of such collaborations are the International Clearinghouse for Birth Defect Monitoring Systems (ICBDMS) and the Surveillance of Congenital Anomalies in Europe (EUROCAT). This approach to environmental surveillance, however, has significant statistical problems and may only detect effects of common environmental factors that are either very unevenly distributed geographically or are introduced suddenly.

BOX 1

Traditional paradigm of birth defect registration and surveillance

- Careful collection of cases in a defined population makes it possible to detect changes in prevalence.
- Once a change in prevalence is detected, its source may be identified through supplementary studies.
- Synchronized surveillance across nations increases the capacity to detect new teratogens.

BOX 2

The historic record of surveillance

- General surveillance gives limited results.
- It is extremely difficult and expensive to operate a reliable registration system.
- Surveillance has a relatively low capacity to detect effects of risk factors (drugs).
- Most birth defect registries must justify their existence by also using data for other (research) purposes.
- Environmental teratogens and clusters are still issues of concern.

There are limited results from general surveillance through registries, although registries played a role in the detection of the two teratogens, thalidomide and valproic acid. A shift in public concern from teratogenic effect of drugs to teratogenic effects of other environmental and occupational exposures still motivates the awareness represented by birth defect registries.

BOX 3

Types of registries used for prevalence estimation

- Population-based birth registries with identification of cases.
- Case registries with good estimates of numerators (number of births in population).
- Case registries with vague estimates of population size.

The Medical Birth Registry of Norway (MBRN) is a population-based registry that covers all births in Norway and has moderately good estimates of birth-prevalence of several birth defects (Lie et al., 1994).

BOX 4

To enrich a registry with exposure information on cases and controls

- Target specific types of exposure (medication, smoking, vitamins, alcohol).
- Collect random controls from a defined population in population-based registries, if possible.
- Consider collecting matched controls (hospital-based) from case-based registries.
- Be aware that it costs more to collect good controls than to collect cases.

An alternative to making environmental surveillance conditional on good prevalence data could be to collect information directly on a set of candidate exposures. Hypothetically, the collection of arbitrary cases with

Information on key exposures – medication, smoking, intake of alcohol and vitamins – is important ...

matched controls that have information on key exposures would shortcut the expensive first step of collection of prevalence data. It would also make it possible to run studies of the effects of several exposures once sufficient amounts of data are collected. To the degree that candidate exposures are identified and may be studied in matched case-control studies, this appears to be an attractive alternative to collection of prevalence data, particularly in populations where prevalence data are hard to collect. However, there are inherent problems with such matched case-control studies. Problems with the quality of the controls, as well as recall bias, may be serious enough to make it difficult to defend a pure case-control approach to surveillance of environmental exposures. Regardless of whether prevalence data are available or not, pooling of case-control data with cohort data on exposures in meta-analyses may serve as another basis for collaboration and comparison (Wacholder et al, 1992a-c).

During a period in the 1990s, the MBRN collected drug exposure information on birth defect cases and matched controls. This was done as part of a collaborative project called “MADRE”, coordinated by ICBDMS. The collection targeted information on medications noted on antenatal records. Such prospective information should be free of recall bias. However, the data appeared to be incomplete and had limited value for most medications. The importance of having information available on key exposures (medication, smoking, intake of alcohol and vitamins) motivated MBRN to redesign the registry and, since 1999, they have collected this information on all births.

BOX 5

Limitations of matching versus random controls in case-control studies

- Exposures that are correlated with matching criteria cannot be studied (geographical environmental exposures, time etc. are difficult to estimate).
- Matched controls do not represent the population (cannot estimate exposure prevalence, attributable risks, gene frequencies, Hardy-Weinberg equilibrium, etc).
- Each case-group need separate controls.
- Matching requires special statistical techniques.

The MBRN has also served as a base for a case-control study of CLP, the cases not being recruited from the registry, but from treatment centres in Norway. For the collection of controls, however, the registry was instrumental. Choosing candidate controls by random selection from the whole population ensured that the controls represented the same population as the cases. There was, however, no matching involved so it was strictly a case-cohort design.

BOX 6

Studies of genes, exposures and GEI

- Case-parent triads may be more informative and even easier to collect than (matched) case-control data.
- (Matched) case-control data will enable estimates of the main effects of the exposures.

There are many variations in the data codes of different registries...

To date, discussions have addressed the possible detection of environmental exposures. Despite their limitations, case-control studies may yet be a supplement to prevalence registries. In some situations where prevalence data are extremely hard to get, case-control data may still serve some surveillance purposes. If the aim is to study genes or GEI, however, traditional case-control data may be less attractive than data and biological samples collected only from the cases and their biological parents (Weinberg et al., 1998; Wilcox et al., 1998). Comparison of such data from different sources should be relatively unproblematic.

1.3 Definitions, classifications and coding of birth defects and craniofacial anomalies

For the purpose of handling and exchanging data, it is customary to translate words into codes. Here again, although most birth defect registries use codes based on the WHO *International Classification of Diseases* (ICD), there are many variations, and some registries use their own codes in preference to those in the ICD. The ICD has to deal with the entire scope of human diseases so it is not sufficiently detailed for many specialized purposes; extensions of its codes have therefore been developed by some organizations. Among the best known are:

- the British Paediatric Association (BPA) extension, which covers all medical conditions of childhood,
- the US Centers for Disease Control and Prevention (CDC) extension of the birth defects section, and
- the extension, for birth defects only, devised by EUROCAT.

At the time of writing this report, the most recent versions of the ICD are the 10th and 11th editions (ICD-10 and ICD-11). In addition to problems with the terminology of birth defects, there are also some difficulties with the epidemiological terminology. After long years of debate, the preferred term to describe the frequency of occurrence of birth defects is “prevalence” rather than “incidence”. Traditionally, defects in liveborn and stillborn infants have been reported in terms of birth prevalence rates,

usually per 10 000 related births. This remains appropriate for defects unaffected by prenatal diagnosis and terminated pregnancies. However, as these techniques are applied to an ever-increasing range of defects in many countries, the term “birth prevalence rate” becomes increasingly inappropriate because terminated pregnancies are not births in the usual sense. By the same token, the term “fetal defect” may be preferable to “birth defect” in these cases. For defects subject to prenatal diagnosis and termination of pregnancy, the term “prevalence” may be sufficient. “Total prevalence” has been suggested, but logically this should include defects in spontaneously aborted fetuses and embryos.

There are
to date
no internationally
accepted terms
for defining
birth defects ...

The increasing practice of terminating pregnancies makes it necessary to consider the appropriate denominator to use in the calculation of rates. The traditional denominator is the total number of live births and stillbirths. If significant numbers of fetuses – most of which would have become either live births or stillbirths – are being treated by terminating the pregnancy, an addition to the denominator seems appropriate. Logically, this should be the total number of pregnancies terminated – for whatever reason – but, because this figure is not always easy to obtain, many registries add the number of pregnancies terminated because of fetal defects to the total live births and stillbirths. As technical advances now allow prenatal diagnosis and termination of pregnancy earlier than in the past, the possibility arises that some of these fetuses would have aborted spontaneously. This could result in an apparent (but not real) increase in prevalence rates. Very early termination of pregnancy also adds to the difficulties of “postnatal” confirmation of prenatally diagnosed defects.

These problems are not likely to make a profound difference to reported rates of congenital anomalies. They are simply matters to be borne in mind when making comparisons between reported rates from different registries.

1.3.1 Terminology

It is unfortunate that there are to date no internationally accepted terms for defining birth defects. Attempts have been made to give specific and restricted meanings to such words as *association*, *deformation*, *malformation*, *sequence*, *syndrome*, etc., but the recommendations published from time to time have not been universally adopted. Even more important than these global terms are the definitions of individual defects. There is scarcely a single defect or group of defects that does not present some extremely difficult problem. For example, several common birth defects are abnormalities of size. Most often, the affected part is abnormally small (microcephaly, microphthalmos, microtia) but it may also be abnormally large (macrocephaly, megalocornea). The measurements upon which these diagnoses are based are continuous variables, but the extent

An array of terms has evolved over several centuries to describe craniofacial and other malformations ...

of the variation from the mean which underlies such diagnoses, in terms of centiles or standard deviations, is rarely defined. Furthermore, the actual measurements are not necessarily very precise or reproducible.

Infants with more than one defect (multiples) present a different problem in classification. If the combination of defects constitutes a recognized syndrome, sequence or association, the appropriate collective term will commonly be used. When the defects do not add up to any recognized condition, a decision must be made as to how to record them. If each defect is recorded separately, the number of defects on record will, of course, exceed the number of affected infants.

There are a few widely adopted, pragmatic conventions. For example, extensive non-closure of the neural tube may result in anencephaly and spina bifida. In these cases, the spinal lesion is regarded as an extension of the cranial lesion and spina bifida is not usually recorded as a separate defect. Defects that are consequential upon other defects (e.g., hydrocephalus, talipes and dislocated hips associated with spina bifida) are often not counted as separate defects.

The same problem may arise in relation to individual organs. Congenital heart disease is frequently complex. If the four components of Fallot's tetralogy are present, the case will be given the single diagnosis, Fallot's tetralogy. However, if there are multiple defects that do not constitute a recognized *syndrome*, a decision must be made on whether to record one defect (in which case, which one?) or all. Also within the field of congenital heart disease are the problems of the ventricular septal defect (which frequently closes spontaneously) and patent ductus arteriosus (which is very common in small, pre-term babies and may require active measures to close it).

In the tenth revision of the *International Classification of Diseases* (ICD-10), the term "congenital anomalies" was replaced by "congenital malformations, deformations and chromosomal abnormalities" to denote structural malformations and exclude conditions such as inborn errors of metabolism. "Craniofacial malformations" is therefore an appropriate term.

An array of terms that is sometimes confusing has evolved over several centuries to describe craniofacial and other malformations. After consultation with professional colleagues and teratology researchers, the Human Malformation Terminology Committee of the International Federation of Teratology Societies (IFTS) has developed a comprehensive list of the various terms used for congenital malformations, deformations and so-called disruptions. The IFTS Committee's list provides "preferred" terms, "acceptable" terms, "non-preferred" terms and definitions.

BOX 7

Examples from the IFTS Committee's terminology list

Preferred term	Acceptable term	Non-preferred term/s
Scaphocephaly	Dolichocephaly	Leptocephaly; Boat head; Dolichocrania Cymbocephaly; Mecistocephaly
Microcephaly	Microcrania	Nanocephaly
Cleft lip		Harelip; Cheiloschisis; Stomoschisis; Stomatoschisis

1.3.2 Working definitions

Craniofacial anomalies (CFA): this term covers a poorly defined group of congenital anomalies named after the anatomical location of a given defect present at birth. According to working definitions, it could include any etiologic category (chromosomal, environmental, Mendelian, multifactorial, etc.), as well as any pathogenetic mechanism (malformation, deformation, disruption, dysplasia), or any clinical category (developmental field complex, isolated defect, sequence, syndrome, etc). Therefore, in this work we will only refer to oral clefts, including typical cleft of the lip and/or palate as an example of CFA. In the future, other congenital anomalies of this group could be considered (see Table 1 below). Whenever not specified otherwise, we will refer in this work to the isolated forms of oral clefts, without other congenital defects detected in the same child.

Oral clefts (OC), including cleft lip, with or without cleft palate (CL/P), and isolated cleft palate (CP), occur in approximately 1 in every 700 live births; that is, with about the same frequency as Down syndrome, neural tube defects, polydactyly, and other so-called “common” congenital anomalies.

Registries and registers, monitoring and surveillance: If the primary objective of the planned system is not to monitor but to *register* oral clefts in the specific sense of the word – to *register*, namely, to recruit and follow up cases in a central repository (Last, 1995) – the quality of recorded data should be of more concern than completeness of ascertainment. (See Box 9 below.)

BOX 8

ICBDMS definitions

The definitions and characteristics summarized here were published by the International Clearinghouse for Birth Defect Monitoring Systems as norms for the 28 participating programmes (ICBDMS, 1991, 2001).

Even though minor deviations can be adopted for other studies, a detailed definition including a list of inclusions and exclusions must be compiled and agreed upon before material from different registries can be collected.

- **“Cleft lip with or without cleft palate (CL/P):** a congenital malformation characterized by partial or complete clefting of the upper lip, with or without clefting of the alveolar ridge or the hard palate. Excludes midline cleft of upper or lower lip and oblique facial fissure (going towards the eye) (ICBDMS, 2001).

“Cleft of the lip arises by non-fusion of various processes that build up the face. Clefting of the lip (“harelip”) may also include the alveolar processes and the palate. Clefts may be unilateral, predominantly on the left side, or bilateral. Most infants with clefts of the lip have no associated malformations. Clefts may be surgically repaired with some functional and cosmetic restoration, but often a series of treatments is necessary until the child is of school age or more. There is a genetic background to cleft lip and a recurrence risk exists for siblings. Exogenous factors probably play a role, e.g. certain drugs and maternal smoking. Facial clefts may be detected prenatally by ultrasound, but usually not until late pregnancy (ICBDMS, 1991).

- **“Cleft palate without cleft lip (CP):** a congenital malformation characterized by a closure defect of the hard and/or soft palate behind the foramen incisivum without cleft lip. Includes sub-mucous cleft palate. Excludes CLP, cleft uvula, functional short palate, and high narrow palate (ICBDMS, 1991).

“This condition is characterized by a cleft throughout the soft and hard palate, usually positioned in the midline, but without clefting of lips or alveolar process. The cleft originates in the non-fusion of the maxillar palatal processes during the tenth week of embryonic development, but details of the pathogenesis are debated. The cleft may be sub-mucous. Some clefts are associated with a small lower jaw (micrognathia) where the tongue may have mechanically prevented fusion of palatal processes (Pierre Robin sequence). CP is often associated with other malformations in various syndromes. Genetic factors play a role in non-syndromic isolated cleft palate, and a recurrence risk exists in siblings. Environmental factors may increase the risk for an isolated cleft palate, e.g., certain drugs and, possibly, maternal smoking. The palatal cleft impairs swallowing and, later, speech. Treatment usually starts with a prosthesis to cover the cleft, followed by surgical repair. If no chromosomal abnormalities or serious malformations are associated, prognosis is excellent (ICBDMS, 1991).”

Table 1: The spectrum of CFA*

Cleft lip (CL)	Cleft lip with or without cleft palate	Cleft lip and/or palate; oral clefts	Craniofacial anomalies
Cleft lip & palate (CLP)			
Cleft palate (CP)	Cleft palate		
Oblique facial cleft		Atypical facial clefts	
Macrostomia			
Midline cleft lip	Holoprosencephaly		
DeMyer facies			
Other CNS defects		CNS defects	
An/microphthalmus		Eye defects	
Other eye defects			
An/microtia		First and second branchial arch defects	
Preauricular tags			
Preauricular sinus			
Agnathia		Mandibular defects	
Micrognathia			
Other jaw defects			
Anodontia		Dentition defects	
Other dental defects			

* This list of CFA is not exhaustive.

BOX 9
Aims of a coordinating registry

In terms of the above concept (*para. 1.3.2*), a coordinating registry should aim to:

- build up a collaborating network, with the participation of all member registries, as a permanent activity, suitable for descriptive epidemiology, surveillance (including monitoring), activities in preventive public health, interactions with support organizations, education and training;
- conduct research programmes, with the participation of some member registries, as temporary, short or long-term activities aimed at specific objectives.

1.4 Clinical support of craniofacial anomaly registration

Three interrelated research issues, within the clinical theme, were addressed at the meeting.

1.4.1 Evidence-based care

Evidence-based care focuses on the replacement of current widespread uncertainty and confusion in clinical care with a sound evidence base derived from rigorous clinical research. There is a pressing need to mobilize a critical mass of clinical research expertise and to access sufficiently large samples of patients for adequately-powered clinical trials. Initial efforts should include:

There is an urgent need for work on psychological and quality-of-life measures ...

- 1) Trials of surgical methods for the repair of different orofacial cleft subtypes, not just unilateral clefts.
- 2) Trials of surgical methods for the correction of velopharyngeal insufficiency.
- 3) Trials of the use of prophylactic ventilation tubes (grommets) for middle-ear disease in patients with cleft palate.
- 4) Trails of adjunctive procedures in cleft care, especially those that place an increased burden on the patient, family or medical services, such as presurgical orthopaedics, primary dentition orthodontics and maxillary protraction.
- 5) Trials of methods for perioperative pain, swelling and infection management, and nursing.
- 6) Trials of methods to optimize feeding before and after surgery.
- 7) Trials addressing the special circumstances of care in low- to middle-income countries in respect of surgical, anaesthetic and nursing care.
- 8) Trials of different modalities of speech therapy, orthodontic treatment and counselling.

Equally urgent is the need to either create collaborative groups or improve the networking of existing groups in order to develop and standardize outcome measures. There is an especially urgent need for work on psychological and quality-of-life measures, and economic outcomes.

For rare interventions, prospective registries should be established to hasten collaborative monitoring and critical appraisal, equivalent to Phase I trials. Relevant topics would be craniosynostosis surgery, ear reconstruction, distraction osteogenesis for hemifacial macrosomia and other skeletal variations, midface surgery in craniofacial dysostosis and correction of hypertelorism.

1.4.2 Quality improvement

Quality improvement requires the development and dissemination of methodologies for monitoring and improving the delivery of clinical services.

The international adoption of a set of guidelines for the provision of clinical services and the maintenance and analysis of minimum clinical records of cleft care is proposed. Various registries of clinical outcomes have recently emerged and are working independently. Efforts should be made to harmonize them.

1.4.3 Access and availability

Access and availability requires the identification of strategies to maximize access to adequate levels of care for all affected individuals, irrespective of nationality. In large parts of the world, routine public health care services are unable to afford treatment for CFA. Three general approaches can be identified:

- high volume indigenous centres of excellence,
- contractual agreements initiated by NGOs with local hospitals, and
- voluntary short-term surgical missions.

The long-term benefit of these efforts could be developed by:

- 1) A survey of the charitable organizations involved and the scale of their work.
- 2) An appraisal of the cost-effectiveness and clinical effectiveness of the different models of aid.
- 3) The promotion of dialogue between different NGOs to develop commonly-agreed codes of practice and the adoption of the most appropriate forms of aid for local circumstances, with an emphasis on support that favours indigenous long-term solutions.
- 4) The initiation of clinical trials concerning the specifics of surgery in low- to middle-income country settings, one-stage operations, optimal late-primary surgery, anaesthesia protocols (e.g., local anaesthetic, inhalation sedation), antisepsis.
- 5) The development of common core protocols for genetic, epidemiological and nutritional studies alongside surgery.