

Guide to the Reader I: What is Presented in the Tables and Maps?

In the tabulations that follow, data on liveborn and stillborn cases of birth defect are presented for 52 registries in five continents, 1993-98, in a total population of 16.9 million births.

For 32 of these registries, data on terminations of pregnancy following prenatal diagnosis are also presented separately. These are registries, which have been able to collect information on such cases, in countries where termination of pregnancy for congenital anomaly is legal.

The data presented relate to a selection of major structural and chromosomal defects. They do not include inborn errors of metabolism, nor do they specifically address disabilities arising from defects of vision, hearing and intellect, although some of the anatomical defects presented are associated with such disabilities.

The tabulations provide information about a wide range of common anomalies. For each malformation, we present two tables, one graph and three maps.

The first table contains (for every program/registry that provided data for that anomaly):

- Years coverage (years covered between 1993 and 1998).
- Cases: number of live + stillborn cases.
- Births: denominator i.e. number of births in each population (live and still).

Birth Prevalence per 10,000: birth prevalence rates per 10,000 births calculated as (live + still born cases) / births with 95% Confidence Interval. We adopt the standard convention of expressing this as “prevalence” rather than “incidence” (Hook, 1982). We present the confidence interval only if a test of heterogeneity calculated within the registry on yearly data was not statistically significant. If significant, the values reported (as bold character) represent the birth prevalence ranges (minimum and maximum values of the period). This is because, if the values are heterogeneous, the confidence interval around the point estimate may be misleading.

L+S trend: for every program, we tested the linearity of the trend in birth prevalence considering only the live and still births time series using the Chi-squared test for trend. When the test is significant the arrow indicates the increasing (upward) or the decreasing (downward) trend.

The second table refers to programs/registries, which reported terminations of pregnancies (ToP). For each year under investigation (1993-1998) we show the total number of cases (live + still births + ToP) and the percentage of ToP (calculated as ToP / (live+still births + ToP)). Note that the time period covered could be different from the first table, as some registries did not have information on terminations of pregnancy for all years. The column “trend” shows the significant trends (tested using the chi-squared test for trends) of the total number of cases.

The graph shows the prevalence at birth per 10,000 with registries ordered by magnitude of point estimate of prevalence, along with the 95 % confidence interval (i.e. a graphical presentation of the data given in the first table). The registers that in the first table evidenced heterogeneity of trends have no bars and the intervals show the range of the birth prevalence.

The maps show the geographical representation of the birth prevalence rates at birth (i.e. the data presented in the first table), shaded by quartiles of the rate distribution.

Guide to the Reader II: Factors affecting the accuracy of estimation of birth prevalence

Many factors affect the accuracy of estimation of prevalence rates at birth. Readers can refer to the following publications for more information: Mastroiacovo & Botto 2000, ICBD 1993, EUROCAT Working Group 2002. Below we give a brief summary of the main factors.

1. Definition of the population.

Registries can be “population-based” or “hospital-based”. “Population-based” means that they cover residents of a defined geographical area. “Hospital-based” means that they cover births in selected hospitals. Where a registry is hospital-based, it is possible that there has been some selection of high-risk pregnancies towards or away from the selected hospitals, and thus estimated prevalence rates may be biased upwards or downwards. If a registry is population-based, it must ensure coverage of residents who deliver outside the geographic boundaries, who may also be at higher or lower risk than the rest of the population. In practice, there are also some variants of the above definitions based on knowledge of how information can be gathered and where mothers go to deliver. Assessing the potential for bias requires detailed knowledge of the local situation. The definition of the population covered by each registry is given in the “Registry Descriptions” section of this Atlas.

2. Definition and classification of cases and diagnostic practice

Epidemiological data are derived from diagnoses made by clinicians working within given health service conditions. A registry is rarely in a position to impose a standard definition or diagnostic practice, though it may facilitate the adoption of standards. Many malformations exhibit a range in severity and the inclusion or exclusion of mild forms may be a source of variation between prevalence rates. This is of greatest consequence for widely variable conditions like hypospadias, where the very mild forms that represent a large proportion of all cases may or may not be counted. Other examples are microphthalmia and microtia. Another source of variation is the definition of stillbirths included in prevalence rates. Although stillbirths are included in the data of all reporting registries, they are defined variably by a minimum of 16, 20, or 28 weeks of gestation or by birth weight limits of at least 500 or 1,000 grams. As malformed infants tend to be born prematurely or to be stillborn, the inclusion or exclusion of stillbirths of low gestational age or weight may lead to significant variations in the reported prevalence of some birth defects. Many variations in diagnostic practice may affect the reported prevalence of birth defects. For example, the accurate reporting of chromosomal anomalies (e.g. Trisomy 13 or 18 and Down Syndrome) is dependant on karyotyping rates and indications for karyotyping. The autopsy rates for stillbirths and neonatal deaths will determine the likelihood that a birth defect is diagnosed, or the accuracy of the diagnosis, especially for conditions which are not externally visible such as serious congenital heart disease (e.g. hypoplastic left heart syndrome).

Children with syndromes and multiple anomalies present particular classification problems. Practice varies as to whether the name of the syndrome only is recorded, or all of the component malformations. Defects that are seen as consequences of other defects (e.g. hydrocephaly when associated with spina bifida) are generally not recorded separately. Such combinations of malformations are called “sequences”.

3. Ascertainment and coding

A diagnosis must not only be made, but also be recorded accurately, with the record reaching the registry, often through one or more intermediary records. There can be a loss of information between the place of diagnosis and the registry, either in terms of whether the child is recorded as having a birth defect at all, or in terms of the detail or accuracy of the diagnosis recorded (e.g. whether the baby is recorded with congenital heart disease, or specifically with coarctation of aorta). Registries work hard to establish and maintain an information pathway which will lead to high case ascertainment (i.e. the proportion of diagnosed cases who are registered), and accurate diagnostic information. Another step where there can be loss of information is between the text diagnostic information and the coding of that information. Most registries use versions 9 or 10 of the International Classification of Disease, some with special extensions (a further one or two digits) to allow more detail to be recorded.

Overall case ascertainment probably never reaches 100%, and its level depends on a registry's methods of data collection. Registries need to use multiple sources of information. Underascertainment of some anomalies can occur if sources of information stop in the early neonatal period, as diagnoses may be made later than this. Specialist services treating children later than the postneonatal period are also vital for confirmation of diagnostic details. Some birth defects are now being discovered earlier in life due to prenatal and postnatal screening programmes. For example, cystic kidneys are more likely to be diagnosed early in life if there is ultrasound screening of the kidneys. This can lead to variation in prevalence rates between regions and over time as screening practice changes.

It is more difficult for registries covering very large populations to attain a high level of case ascertainment, although some large registries are organised hierarchically with local offices. Local contact with clinicians and other information sources is vital. The level of resources available to the registry or local office to employ suitable personnel, and the stability of those resources to retain experienced personnel, will also affect the quality of the data collected. Data management itself can be complex, even the conceptually simple tasks such as not registering the same child twice.

4. Statistical Considerations

A 95% confidence interval is an estimated range of values with a 95% probability of covering the true population value. When the number of cases on which the prevalence rate is based is small, the confidence interval is wide, representing a large degree of sampling error.

Statistical significance (meaning a low p value below an arbitrary threshold such as 5%) represents how likely differences in prevalence could have arisen by chance. Unfortunately, p values and statistical significance are often accorded too much weight. Critical readers should bear in mind that the $p < 0.05$ threshold is wholly arbitrary. To call one finding significant when the p value is 0.04 and another not significant when it is 0.06 vastly overstates the difference between the two findings, moreover 10% or 1% significance levels could be reasonable alternatives. Because p values are quantifiable and seemingly objective, it is easy to overemphasize the importance of statistical significance. For most studies, the biggest threat to an author's conclusion is not random error (chance), but systematic error (bias). Thus, readers must focus on the more difficult, qualitative questions discussed in this Guide. When using numerous statistical tests with a significance level of 5%, one can expect 5% of test results to be spuriously significant, i.e. the differences in prevalence are due to chance alone. As many

statistical tests were performed for this Atlas, this needs to be borne in mind. We have presented in this Atlas straightforward maps where prevalence rates have been divided into quartiles of the rate distribution. Readers should be aware of the interpretational “traps” of this method. Firstly, the eye tends to focus on the large areas in the map of the same colour, rather than taking into account smaller areas which may be much more densely populated. This can lead the eye to see false geographical patterns. Secondly, some of the rates are based on quite small numbers of cases with wide confidence intervals around the prevalence rates (as shown in the first table for each birth defect). The map does not show the confidence interval, only the point estimate, and this can lead to misinterpretation of apparent patterns. Small populations moreover tend to give rise to the most extreme (high or low) rates due to Poisson (chance) variability, even if the true disease rates are similar across the areas, focusing viewers' attention on these extreme areas when they scrutinize the map. Bayesian smoothing has been employed as a method to get round this problem in mapping (Bernardinelli, 1995; Osnes, 1999). In this Atlas however, our use of maps is as a simple graphical representation of data given in tables, rather than as a tool for presenting the strength of evidence for geographical variation, which would require more attention to the range of factors affecting accuracy of estimation of prevalence rates summarized above in this Guide.

Guide to Reader III: Terminations of pregnancy: a challenge for birth defect epidemiology

Terminations of pregnancy are a special challenge to birth defect registries and to the whole field of birth defect epidemiology. We know that prenatal screening policies (and the resources for prenatal screening) vary enormously between different countries and between regions and even hospitals within countries. The “culture” in terms of how often prenatal diagnosis of a birth defect leads to termination of pregnancy also varies. For example, termination of pregnancy is very widespread for lethal conditions such as anencephaly, but the practice is much more variable for conditions such as spina bifida. Thus, prenatal screening followed by termination of pregnancy introduces considerable geographic and temporal variation in prevalence rates at birth, and the proportion of terminations must be known or well estimated to assess whether there are real differences in “risk” between populations related to genetic or environmental risk factors.

Many birth defect registries in Europe have been set up to provide a mechanism for the audit of prenatal screening practice. The registry can provide data on the proportion of cases of congenital anomaly diagnosed prenatally, the proportion of positive prenatal screening results which were confirmed as cases of congenital anomaly, and the proportion of prenatally diagnosed cases which led to termination of pregnancy, as well as related information about prenatal screening methods.

Registries often require access to entirely different sources of information to ascertain terminations. Assessment of completeness of ascertainment of terminations requires detailed knowledge about local use of services (public and private) and information flows.

Ideally for epidemiologic purposes, terminations of pregnancy should be subject to the same rigour of diagnostic verification as live and stillbirths, but this is not always so. For example, autopsies may not be carried out to confirm the diagnosis, and a karyotype may not be performed where multiple malformations have been detected prenatally by ultrasound, to determine whether a chromosomal anomaly is present.

Prenatal diagnosis presents particular problems for hospital based registries (see definition of population above) since it is common for referral to a tertiary centre of expertise to take place, either for termination or for the birth of the affected child. This may increase the potential for bias in the estimation of prevalence or in the estimated proportion of terminations.

Reporting of terminations of pregnancy can lead to relative “overascertainment” of cases. The earlier in pregnancy the termination, the greater the probability that the pregnancy would in other circumstances have ended naturally in a spontaneous abortion. A spontaneous abortion would not necessarily have been examined for malformations or reported to the registry. These probabilities are generally small, but when the number of early terminations are high might result in a slight inflation of the total number of cases recorded compared to what would be expected if no terminations had been performed.

Prenatal screening and diagnosis, whether or not followed by termination, can also lead to relative “overascertainment” of cases when the age of detection of a congenital anomaly is brought within the age coverage of the registry. This obviously depends on the age limit each registry applies to its information gathering, as well as diagnostic practice regarding the age when the anomaly would usually be detected postnatally. Similarly, the recorded proportion of

all cases which are terminations of pregnancy may be inflated if prenatally diagnosed cases are ascertained by the registry more completely than postnatally diagnosed cases.

The purpose of prenatal diagnosis is to increase the possibility of optimal management of the pregnancy and baby. While the issues of prenatal diagnosis and termination of pregnancy are intertwined in the evaluation of prevalence rates based on epidemiologic data, they are not intertwined in health service terms. Prenatal diagnosis can lead to beneficial outcomes such as effective early neonatal treatment or care. As outcomes improve, the practice of termination may well change.

Birth defect registries are concerned to provide a basis through surveillance and research for the primary prevention of birth defects. Currently, we can expect the prevalence of neural tube defects at birth to decrease through the combined impact of primary prevention and prenatal screening and termination, and the challenge is to vastly increase the proportion of cases prevented by appropriate folic acid supplementation (see Introduction). The larger challenge is to identify the range of genetic and environmental risk factors for birth defects in order to increase the potential for primary prevention.

References

Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M Bayesian-analysis of space-time variation in disease risk. *Stat Med.* 1995 Nov 15-30;14(21-22):2433-43.

EUROCAT Working Group. EUROCAT Report 8, University of Ulster, 2002.

Hook EB Incidence and prevalence as measures of the frequency of birth defects. *Am J Epidemiol.* 1982 Nov;116(5):743-7.

Mastroiacovo P, Botto LD Surveillance for Birth Defects and Genetic Diseases. In: *Genetics and Public Health in the 21st Century*. Eds. MJ Khoury, W Burke and EJ Thomson, Oxford University Press, New York, 2000.

Osnes K, Aalen OO Spatial smoothing of cancer survival: a bayesian approach. *Stat Med.* 1999 Aug 30; 18(16): 2087-99.

The International Centre for Birth Defects. Guidelines for the development on national programmes for monitoring birth defects. World Health Organization, Hereditary Diseases Programme. Geneva, 1993.

