MONITORING HUMAN BIRTH DEFECTS: METHODS AND STRATEGIES

ERNEST B. HOOK
NEW YORK STATE DEPARTMENT OF HEALTH
and
ALBANY MEDICAL COLLEGE OF UNION UNIVERSITY

1. Introduction

Since the rationale and goals of a paper in multidisciplinary settings are often unclear to those outside the author's specialty, a statement of purpose may not be out of place. My interest is in the diminution of birth defects and detection of preventable environmental causes of such events.

This paper has been written assuming the reader has no previous knowledge of birth defects. Definition of medical terms not provided in the text are not crucial to the argument, but may be found in any medical dictionary. The record of the New York State Birth Defects Institute's Symposium of October, 1970 will present in much greater detail some of the themes discussed here [10].

Some of the material treated here was originally presented in more condensed form at a National Foundation Symposium in New York City on Environment and Birth Defects, January 27, 1971.

2. Definition and incidence of birth defects

A major human birth defect may be defined as an anatomical structural variant that produces a significant clinical or cosmetic effect. This definition is, of course, a somewhat loose one in that what may be abnormal in one setting may be acceptable in another. For the purposes of monitoring as discussed here, this vagueness will not be a problem. But it should be pointed out that reports of incidence of total birth defects in various groups cannot be compared unless the precise defects scored by the authors and their method of ascertainment are specified [11]. For the purpose of this discussion, the incidence of infants with major defect detectable at birth will be assumed to be about two per cent. The incidence of particular malformations is of course much rarer. Order of magnitude estimates for some of the most frequent major malformations are (in contemporary U.S.A.), anencephaly $1 - 2 \times 10^{-3}$, mongolism $10^{-4}$, spina bifida $0.5 - 2.0 \times 10^{-3}$, polydactyly (whites) $0.4 - 0.7 \times 10^{-3}$, polydactyly (blacks) $5 - 10 \times 10^{-4}$. Some might call the latter a minor malformation. (The definition of a minor birth defect is given in Section 8.) But most types of major
malformations are probably very rare in liveborns having an incidence of no greater than $3 - 5 \times 10^{-6}$.

3. Defects by likelihood of ascertainment

For the purposes of ascertainment we can divide defects by their likelihood to be diagnosed at birth. (a) Externally dramatic defects include such malformations as anencephaly, cleft lip, and phocomelia. (b) Externally detectable malformations likely to be noted in superficial physical examination at birth but not likely to be starting on initial observation include such defects as syndactyly and imperforate anus. (c) Cryptic but diagnosable defects include among others, structural abnormalities of the genitourinary system (for example, polycystic kidney) and some congenital heart disease. (d) Cryptic defects not likely to be diagnosed or diagnosable in infancy may include occult tumors, Merkel’s diverticulum, and so forth. Some in this category may only be found at autopsy.

4. Defects by causal mechanism

At least five classes may be distinguished, (a) defects genetically inherited such as simple Mendelian recessive or dominant traits, (b) defects associated with inherited chromosomal translocations (which may be included with some of class (a), (c) defects occurring as a result of a germinal mutation (in the previous generation) resulting in the appearance of either a simple Mendelian trait or a fresh chromosome abnormality, (d) defects produced by some agent to which the mother is exposed during gestation, and (e) defects produced by some unknown concatenation of environmental and genetic events, each of presumably relatively small effect but which interact to produce a defect (the multifactorial hypothesis). These categories are obviously not mutually exclusive. Among other things, the effect of any environmental agent may be modified by the genetic background and the effects of genetic factors may be influenced by environmental events during or before pregnancy. This has been well documented in nonhuman species and probably accounts for at least some of the fluctuation in expression of known single genetic or environmental causes of human birth defects. Thus even the boundaries between these categories of causes are not sharp.

From the viewpoint of the environment, an external factor may produce defects by a number of different mechanisms. It would be hazardous to suggest all possible routes by which structural development can go awry. The list below is of only very general categories but may still be incomplete. An agent may induce defects by (1) inducing a chromosomal translocation or chromosomal nondisjunction or a simple point mutation in germinal cells which then eventually results in a gamete carrying an abnormal genotype; (2) producing somatic mutations in cells of the developing fetus; (3) inducing structural disorganizational events during gestation, directly, or (4) affecting the mother’s physiological
metabolic pathways in such a manner that the effects are dysmorphogenetic in the fetus she is or will be carrying. Agents producing the first two types of events may be said to be mutagens, those producing the latter two types, teratogens. The effect of a known teratogen varies markedly with the time of exposure of the fetus. For most teratogens there is a critical period, usually relatively early in gestation, when the most dramatic effects are produced. In humans, this is almost always at some time in the first two or three months. The literature seems fairly consistent in suggesting that any agent which is mutagenic is also teratogenic. That is, any agent which will induce changes in the genetic constitution of cells will also be dysmorphogenetic if administered at the right time. This may be because accumulation of sufficient mutations during (early) development may eventually result in structural abnormality. The converse however, is not true. Thus, to my knowledge, rubella and thalidomide have not yet been shown to be mutagenic in humans. In any event it seems likely that the magnitude of teratogenic effect is much greater than any mutagenic effect they might possess. In the discussion below we will use the term teratogen to denote any environmental agent that may produce a birth defect. Since toxic effects of environmental agents upon the fetus are not necessarily limited to production of malformation (see Section 8), the term embryotoxin is used for an agent with any deleterious effect upon the conceptus [21].

From the viewpoint of the defect, any single malformation could be produced by any combination of mechanisms mentioned above. But given a single malformation in a particular infant it is likely that no single causal event can be identified, so such occurrences are usually attributed to category (e), the multifactorial hypothesis. A very rough estimate would be that perhaps 90 to 95 per cent of all human defects fall into this latter category which is simply a reflection of our ignorance about the exact pathogenetic events resulting in malformation in such cases. Some may be caused by relatively few “strong” environmental causes. An abrupt increase in a particular defect could be simply a chance event due to concatenation of background events or result from introduction or increase of some environmental insult. (See Section 9 for further discussion of this point.)

5. Detecting causes—the rationale for monitoring

The best way to detect environmental causes of birth defects would be to find some experimental model in which all extraneous background events could be manipulated and controlled and test specific agents. With the possible exception of higher primates there are, however, no such models generally applicable to the human situation. Recently it has been suggested that with liberalization of abortion laws, women about to undergo an elective abortion might volunteer to take a teratogen experimentally. While many factors will be uncontrolled here and a large number of social and ethical questions are unresolved, there is the additional problem that having taken a teratogen, a woman might then change her mind about the abortion. Many obstetricians are thus reluctant to wait more than a brief period after exposure, which may not be long enough for an effect
to manifest. Barring such approaches resort must be made to large scale studies comparing outcomes of pregnancies where the extent of maternal exposure to an agent in question is known. But this is difficult since the exposed and unexposed groups are unlikely ever to be completely identical with regard to other factors. Nevertheless, a good deal of information may be gained thereby, particularly if the effect of the suspected agent is relatively strong compared to the background events.

All the above approaches assume that we have already been able to identify an agent as suspicious on one grounds or another. Of course an optimal social strategy might require testing all compounds and agents to which we are exposed no matter how apparently innocuous. The scale of such an attempt, even given a suitable animal model, is such that at least in the initial approach we would have to limit ourselves to the most ubiquitous compounds and/or those which were already highly suspicious for other reasons. On the other hand by monitoring the incidence of birth defects we hope to have some way of quickly detecting the effect of still unknown teratogens that have, presumably, been recently introduced in the environment (or a recent increase in exposure to already present teratogens). This approach does not necessarily help us to identify the causes of the background defects produced by already existing factors. Nevertheless, since it may help prevent the rate of defects from increasing, it seems at least one reasonable strategy if it is not too demanding, can provide information of interest and possible utility, and is not unduly costly. But it must be remembered that such an attempt is close to being research without hypothesis and subject to all the perils of such an attempt.

6. Aspects of a useful monitoring scheme

These are fairly obvious but are worth specifying. (a) We should have confidence both in the accuracy of the diagnosis of markers used and in our ability to ascertain a high proportion. (b) We should ascertain our markers as close as possible in time to the presumed environmental causal event. (c) The more frequent the background rate of our marker the more likely we are to detect a significant increase in the rate. (d) The larger the population base the greater the absolute number of markers that can be detected and the better the opportunity to study rare defects. (e) The more intense scrutiny we can give to a population, the more markers that can be ascertained. (f) Given the opportunity to monitor rates of markers, we must also have the opportunity to correlate our observations with the environmental exposure of those in whom the marker is detected. (g) The cheaper the scheme the better.

7. Direct monitoring of major malformations

Complete ascertainment of all birth defects is probably impractical. The cryptic category of birth defects may actually be more frequent than those externally
evident at birth. One study showed that about two per cent of all infants at birth had detectable significant malformations [13]. But by age one year perhaps two or three times this number will have had at least one defect diagnosed [14]. Thus it is clear that a monitoring scheme for malformations must essentially limit itself to defects that are likely to be readily detectable. But it should not be overlooked that we still may be missing a significant proportion of malformations in so doing.

One approach to monitoring defects is to use information provided on birth certificates. Almost all states in this country currently request those filling out the certificates (not always the physician delivering the baby) to note whether or not a malformation is present. Thus, there is at least a partial handle on surveillance in the data collected by many health departments.

There are however at least six drawbacks to the use of such certificates for monitoring. (a) As already noted only the externally detectable malformations are likely to be recorded by the individual filling out the birth certificates. (b) Of those diagnosable by gross inspection at birth, a large proportion are overlooked or underreported by the physician or whoever else fills out the birth certificate. As might be suspected the externally dramatic defects, for example, anencephaly and cleft lip which are the most frequent in this category, are relatively accurately reported. The externally detectable less dramatic defects such as imperforate anus or syndactyly are not recorded as accurately and thus the data on such are less useful [5]. Occasionally over compulsive physicians may report anatomical variants that are likely not to be classified as defects by other physicians [15]. A single such individual may create a pseudoepidemic that may be difficult to track down. Furthermore, the particular psychological set of those noting defects in any reporting system may change with time, particularly if recent publicity has called attention to the medical significance of defects that might otherwise have gone overlooked. It is suggested that such an effect occurred with a pseudoepidemic of congenital hip dislocation in England [7]. (c) Diagnostic reports are particularly inaccurate when multiple malformations are present in a baby. The more malformations that are present in an individual the less likely any particular malformation is to be recorded on the certificate. It is likely that only the most externally dramatic ones will appear. Occasionally "multiple malformations" may be noted without any further specification. (d) Even if the birth certificate is filled out as completely and accurately as possible, the coding of the defects that are reported upon such certificates is incomplete in many jurisdictions. Thus, in some instances if two or three defects are present only what may be regarded as the most serious will be coded. (This raises another point concerning clustering of major defects in individuals. An individual with a single specific major defect is more likely than an individual chosen randomly to have another major defect elsewhere.) Another problem with coding is that defects are often grouped together in the rubric of so called "other miscellaneous defects" particularly if they are rare and apparently of lesser significance. Thus a ten fold increase in a particular rare defect may be masked if
hidden in a large miscellaneous category. (e) There is an inertia in any reporting system working from vital records that probably introduces a two month delay in pulling together all the relevant information. The interval between environmental insult and detection of a defect using birth certificates may thus be at least over a year. (f) As already noted the total incidence of external defects visible at birth is about two per cent but the incidence of particular defects is much less frequent. Teratogens may affect only one or just a few organ systems so one cannot lump together all defects but must separate malformations meticulously. Since the incidence of most particular malformations is quite rare and even the most frequent serious external defects have an incidence of no more than one to two per thousand, establishing a doubling or tripling in such rates will be very difficult without an enormous population base.

Nevertheless, despite the difficulties these approaches are already in use in some areas and can be progressively refined perhaps with relatively little time and money. For instance one can include questions on birth certificates relating to specific malformations, thus increasing the likelihood of ascertainment. Secondly, such records can be supplemented by using other sources as well. Large data files exist in many jurisdictions for purposes of justifying hospital insurance payments, characterizing medical care, and for other purposes. These may also provide useful information concerning the incidence of birth defects. Infants with malformations are much more likely, of course, to be hospitalized during the first year of life and to undergo extensive diagnostic evaluation and corrective surgery. A certain proportion will, unfortunately, miss these benefits and some may even die before receiving medical attention. For the latter, information from death certificates may still be of some value. By using such supplemental records one is considering information derived well after birth and thus is even further from the time of the presumed environmental event. The ascertainment with such a system however, is likely to be much better than using birth certificates alone. For instance, Banister in Ottawa using medical insurance records increased the ascertainment of all major defects by three fold up to 3.8 per cent, and similarly Gittelsohn using computerized hospital abstracts had a similar magnitude of effect, compared to that observed using birth certificates alone [1], [6].

One problem that presents itself in the use of such data, as well as birth certificates, is having identified an increase in malformation rate how does one then track down the presumed environmental cause of the increase? The computerized hospital abstracts used by Gittelsohn could be matched with other hospital records but they were still anonymous. At some point someone will have to contact the original family involved to inquire about exposure (unless the records themselves already have such complete data, which is unlikely). The records used in this context have been acquired for other reasons. Individuals who have access to them are thus extraordinarily sensitive to the question of confidentiality, with good reason, but just this problem alone may interfere with attempts to investigate etiology. At least the report of an increase of some defect in a jurisdiction, however, would lead to focusing of attention upon this particular
malformation and would probably lead many individuals in particular regions to look for possible environmental events, without investigation of the central monitoring agency.

While birth certificate data is quite incomplete, newborn hospital records concerning the externally dramatic and externally detectable defects are usually fairly accurate in medically competent settings. Most babies stay in the hospital three or four days before discharge. By that time a number of significant defects are likely to be picked up that were not noted on the birth certificate. Simply by visiting nurseries and delivery rooms regularly in a circumscribed area in which there are a number of hospitals one can get fairly good counts of at least the major defects diagnosed neonatally. In fact, in Atlanta the CDC is currently doing just this [3]. But there is no requirement that an outside agency do the monitoring. Hospitals at teaching centers in large metropolitan areas could cooperate to report monthly the incidence of defects detected on their newborn services.

Information on the incidence of birth defects can also be derived from study of fetal wastage. The proportion of malformations in human spontaneous abortuses is at least twice as high as in human newborns [16], [17], [18]. Thus, the changes in the incidence of malformations here may be helpful in a monitoring scheme. This would require that in particular areas one would have to organize centers for the collection and the examination of fetal wastage. The difficulties with this approach to monitoring are discussed in the next section.

Since teratogens vary widely in the spectrum of effect, at least as far as major malformation is concerned, one cannot monitor for just a few major defects in the hope that the effects of all possible teratogens will be detected thereby. Mutagens, however, are unlikely to have such specific effects in that they are unlikely to have a strong predilection for a particular gene locus. (For the purposes of monitoring mutations, this statement is probably acceptable in that if an agent induces a mutation at one gene locus it will probably do so at many other loci. But differential base pair composition or base sequences in loci may affect DNA reactivity with base analogues or other mutagens, and chromosomal proteins and RNA in one region of a chromosome may also be more likely to bind some specific mutagen.) Since some particular inherited defects (for example, achondroplasia and Apert’s syndrome) are likely to reflect recent dominant mutations, monitoring for these particular defects may provide an index of what is happening to the mutation rate at some other loci. But there are a number of problems with this approach. While a trained diagnostician could probably accurately confirm the diagnosis in suspected cases it would be perilous to rely simply upon outside reporting of these defects. Since the incidence of such defects may be only one in 50,000 births just a few inaccurate diagnoses could produce a pseudoepidemic of mutations. However, as Smith has suggested, all reported instances of such cases could be sought out by knowledgeable diagnosticians in a particular jurisdiction and at least confirmed [19].
It is occasionally stated that the presumed environmental events responsible for a rise in mutation rate are not likely to have a strong effect on the total number of birth defects, but rather upon the genetic load of the population, the effects of which may not be manifest for some generations. Assuming that mutagens also are teratogens however, this may not necessarily be the case.

8. Indirect monitoring of major malformations

If agents which induce major defects produce other effects as well, then monitoring for these secondary or indirect markers may also be helpful. There seems to be good evidence that in the rhesus monkey, a much better model for the human than the rodent, the main pattern of response to an embryotoxin is not the induction of malformation in a liveborn, but production of abortion [21]. This makes biological "sense", in that if a fetus has a defect that makes survival to the time of reproduction unlikely it is considerably more efficient for the species that this organism be lost at the very beginning of (fetal) life. It seems at least plausible to suppose that strong selective factors have produced this as a general category of response to many environmental fetal insults. (One can also imagine why such response might not be as strong in other species with many animals in a litter and greater relative fecundity.) Of course whether abortion or a defect is produced may heavily depend on the dose and the timing of the embryotoxin. If the same pattern occurs in man as well as monkey it is probable that the birth defects we observe reflect only the "upper tip of the iceberg" of human embryotoxicity. The first and the most ubiquitous effects of compounds which have teratogenic effect might thus be fetal loss. It is interesting that the drug amethopterin which is known to be teratogenic in humans, was once used specifically as an pharmacologic abortifacient. The rare cases of amethopterin embryopathy in liveborns may represent only a small fraction of the exposed aborted fetuses.

The outstanding problem with monitoring the rate of fetal wastage however, is that the denominator from which the samples are drawn is simply unknown. (This is also a problem with monitoring the proportion of fetal wastage with detectable abnormality.) Without monitoring the entire female population of reproductive age one cannot tell how many are pregnant at any time and how many have lost a fetus very early in pregnancy without reporting it to a physician. The rate of fetal wastage seems likely to be as high as 25 per cent of all conceptions. But in a unit devoted to collecting fetal wastage, even in the most cooperative hospital the proportion of specimens was no more than six per cent of all live births [17], [18].

Another type of indirect marker that is at least worth considering is low birth weight. One must distinguish two categories: (a) that occurring because of a shortened gestation (SG), and (b) that occurring because of intrauterine growth retardation (IUGR). Of course IUGR and SG may occur together in which case the birth weight will be even lower than that appropriate for the shortened length
of gestation. IUGR has been noted in association with many teratogens. Monitoring the incidence of IUGR requires that data on gestation length as well as birth weight be available. This is usually calculated from the first day of the last menstrual period (LMP) of the mother and the birth date. But often bleeding in the first month or so of pregnancy may be confused with a menstrual period so that the length of time may not be estimated correctly. Frequently the obstetrician makes an estimate of the LMP (or the expected date of confinement, EDC) based upon the size of the fetus, which is significantly different from date the mother recalls. But my suspicion is that the mother is more likely to be correct because an obstetrician is judging the size of the fetus against some normal standard, and if IUGR is present he may underestimate the length of gestation.

Data on the length of gestation in most hospital charts or birth certificates is useless since many medical personnel routinely write "term" or "40 weeks" for all infants not obviously pre- or postmature. (The assumption is that any discrepancy between the interval between LMP and the birth date and "40 weeks" is due to maternal error in memory.) Usually the LMP as stated by the mother is relatively accurately recorded however. In those jurisdictions where "LMP" and not "gestation length" is recorded on birth certificates, one has a handle on monitoring IUGR about as accurate as the monitoring of dramatic defects with vital records. I suspect that such data may be among the most useful that can be derived from vital records in monitoring for environmental hazards. It should be relatively easy to derive these data and furthermore one can measure the extent of IUGR as a continuous variable and thus derive additional power from such an analysis. (The actual variable analyzed would be the "percent of mean appropriate weight for gestational age.") One last point should be made. By timing the length of gestation from the LMP one is adding on two weeks, since ovulation does not occur until about two weeks later (usually). Thus the true length of human gestation is closer to 38 weeks, although most of the medical literature refers to it as 40 weeks.

Another indirect approach to seeking environmental agents that produce birth defects is to use anatomical variants which are of no significance in themselves but which are still correlated with the presence of major defects. Such variants may be called minor birth defects. Relatively trivial abnormalities in flexion creases of the palm, dermatoglyphic ridge patterns, and anatomical shape of the external ear and the eye are included. None of these have any significance per se but analysis has shown some to occur often in association with major malformation or with syndromes of malformations that are produced by the few human teratogens which have been discovered. Obtaining this type information however will require that infants receive systematic physical examinations by trained observers who use precise diagnostic criteria. The main advantage of using such minor defects is that they are very common compared to major malformations. The incidence of all minor malformations taken together varies from perhaps 15 to 30 per cent of the newborn population depending upon what variants are taken to be minor defects, the defects scored, and the population
examined. And the incidence of one particular minor defect is over five per cent in at least one population.

It is still uncertain however, whether minor malformations are likely to be produced by significant teratogens in the absence of major malformations. Thus it is unknown whether an increase in minor malformations might be taken as evidence for a smaller increase in significant cryptic defects in a particular population. Indirect evidence suggests this may be likely but there is not yet definitive proof [8], [9], [13], [19].

9. Strategies in study of causes

It seems unlikely that any particular method discussed above could be recommended to the exclusion of any other. On the other hand, many methods may be used simultaneously in the same setting to provide additional information not necessarily gained from a single approach.

The use of data acquired in surveillance deserves comment. First of all it will provide information on the incidence of particular human major malformations. Data on the indirect markers are also of intrinsic interest. Even if there should be no change at all over a period of time this information is important and it is worth trying to get, if the attempt is not too expensive. Certainly in a negative sense it would be reassuring to know that rates were not increasing. But if rates were found to be increasing, one cannot specify in advance the exact approach to tracking down the cause.

It should be mentioned in this context that extensive epidemiological data are already available on a number of major malformations. Anencephaly for instance, has shown remarkable association with a variety of demographic factors in the population such as low socioeconomic status, elevated maternal age, and so forth [4]. There is a cline in the British Isles that is striking and may even be related to softness of water [20]. Use of old hospital records has revealed that an epidemic of this and related malformations occurred in New England earlier in this century and then subsided [12]. But despite all the data accumulated we still do not yet know what specific environmental factors are responsible for these variations in incidence. Although it seems likely that genetic factors contribute at least something to differences between ethnic groups, it seems unlikely they account for all of it.

Many studies of this and other defects have involved comparisons of "background" rates in different populations. In investigation of the causes of such differences every single variable that differs between the two populations must be considered as possibly significant. On the other hand when following the same population (that is, monitoring) and observing an abrupt increase, one is in a better strategic position because one has at least a particular time to focus upon and a specific population in which to look for the introduction of a new cause. While investigation may be no more productive here it may be easier to develop and investigate plausible hypotheses. In fact when abrupt increases of birth
defects have been observed in particular jurisdictions even practicing physicians have been able to suggest the likely cause. It was in such an anecdotal way that rubella was first suspected as a cause of cataract. With refined statistical techniques in use for systematic monitoring, hopefully we will not have to depend upon personal anecdotal impressions to document a future increase or detect its cause. (But I suspect that inspired guesses have contributed more to our knowledge of single major environmental causes of human birth defects than any methodical discipline.)

Given the presumed "background" causes of major defects a certain number of clusters in time and space are to be expected on probabilistic grounds alone. Thus any single observed rise in a relatively small population may only reflect coincidental occurrence of many "background" causes together. While investigation may disclose no new environmental insult as a likely cause, it may reveal some of the "background" causes. But if there is an increase in the number of clusters in time and space, and if we presume that the increase is caused by some newly introduced environmental factor, we will not of course be able to say a priori which observed cluster is due to the environmental event and which to a concordance of background events, or even sort out such factors in any particular setting. Thus it is clear that all such clusters should be investigated but that we should not be discouraged if unable to identify a new environmental cause in any particular case.

Lastly, there is one point which is perhaps obvious but which is often overlooked. This is that the greatest collection of talent and knowledge from different fields that we can bring to bear on problems of this sort the more likely we will be to reach socially useful conclusions.

REFERENCES