LEADING ARTICLE

Drug Safety 2002; 25 (1): 21-32
0114-5916/02/0001-0021/$25.00/0

How Do We Best Detect Toxic Effects of Drugs Taken During Pregnancy?

A EuroMap Paper

Jørn Olsen,[1] *Andrew Czeizel*,[2] *Henrik Toft Sørensen*,[1,3] *Gunnar Lauge Nielsen*,[4] *Lolkje T.W. de Jong van den Berg*,[5] *Lorentz M. Irgens*,[6] *Charlotte Olesen*,[1] *Lars Pedersen*,[1] *Helle Larsen*,[4] *Rolv T. Lie*,[6] *Corinne S. de Vries*[7] and *Ulf Bergman*[8]

1 The Danish Epidemiology Science Centre at the Department of Epidemiology and Social Medicine, University of Aarhus, Aarhus, Denmark
2 Foundation for the Community Control of Hereditary Diseases, Budapest, Hungary
3 Department of Clinical Epidemiology, University of Aarhus, Aarhus, Denmark
4 Department of Gynaecology and Obstetrics, Aalborg Hospital, Aalborg, Denmark
5 Social Pharmacy and Pharmacoepidemiology, University Centre of Pharmacy, Groningen, The Netherlands
6 The Medical Birth Registry of Norway, University of Bergen, Haukeland Hospital, Bergen, Norway
7 Department of Pharmacoepidemiology & Public Health, European Institute of Health and Medical Sciences, University of Surrey, Surrey, UK
8 Department of Clinical Pharmacology, Karolinska Institute, Huddinge University Hospital, Hoddinge, Sweden

Abstract

It is a major clinical and public health problem that there is no clear strategy as to how we best make use of information obtained when pregnant women take drugs. For this reason, some pregnant women are not treated as they should be and some are given drugs they should not use. We suggest a monitoring system that combines some of the available datasets in Europe. Using these sources as a starting point, one can develop a system that has sufficient power to detect even rare diseases like congenital malformations and sufficient diversity to detect several possible outcomes from spontaneous abortions to childhood disorders. We also suggest that case-crossover designs should be used in case-control monitoring systems that carry a high risk of recall bias. These considerations are based upon our results from a European Union-funded concerted action called EuroMaP (Medicine and Pregnancy).

Drugs are taken by more than half of all pregnant women, and if vitamins and other dietary supplements are included, almost all pregnant women will be users of medication.[1,2] Whether too many or too few are treated is not known, but most likely some pregnant women are not treated the way they should be and some are treated with drugs they should not use. A small fraction of medicines and supplements used during pregnancy probably do have unknown adverse effects that outweigh their therapeutic effects. There is no perfect way to identify these medicines and supplements, but there are several options. This paper discusses some of these options.

Since drugs are not tested in pregnant women before they are released on the market, they will only gradually be used during pregnancy. When they are used, their effects on reproductive outcomes are not reported or utilised in any systematic way in most countries. It is also unclear how to make best use of such findings. In our experience from a European Biomed 2 concerted action called EuroMAP (Medicine and Pregnancy), several data sources are needed and the data collection and analysis have to be coordinated.

Although most classification systems for drugs aim at using all available published information, these systems are constrained by the quality and quantity of information available. This paper discusses how to improve the evidence used in these classification systems – it is high time to prioritise this part of pharmacovigilance in European public health.

1. Background

Because of rapid cell growth and extremely complicated cell differentiation, a fetus is much more vulnerable to drug adverse effects than a neonate, child or adult, and many drugs cross the placental barrier. Exposures that are not very toxic for adults, such as thalidomide or high doses of retinol (vitamin A), may cause serious damage to a fetus. Chemicals that enter the fetal microenvironment may lead to fetal death, congenital abnormalities, functional disorders, reduced fetal growth or a change in organ programming, which may influence susceptibility to diseases later in life. It has been known since 1971 that prenatal drug exposure may be carcinogenic. Diethylstilbestrol given during pregnancy has caused clear-cell adenocarcinoma of the vagina in young women with a predisposition to this disease .[3]

As a result of the thalidomide disaster,[4] most attention has been given to the possible teratogenic effects of drugs since these may have serious consequences for the child, the family and society, and no long-term follow-up is needed for detection. Monitoring of congenital abnormalities is important and we know that some drugs are teratogenic,[5,6] but information is also needed on other health outcomes. Functional defects or diseases that surface to clinical detection later in life should be part of a monitoring system. Several researchers have shown strong associations between indicators of fetal growth and subsequent chronic diseases, which has sparked an interest in adult diseases with a possible fetal origin,[7] but it is not known if drugs play a role in this 'programming'. However, diethylstilbestrol and possibly also some progesterones given during pregnancy may have long-term consequences for the offspring.[8]

In principle, a monitoring system must at least include data on drug utilisation and health outcomes and should be able to detect toxic effects that manifest themselves during pregnancy or shortly after. It should also permit long-term follow-up of selected exposure groups, especially for exposures that could interfere with fetal growth and brain development. The monitoring system should cover drugs used shortly before conception if they are slowly metabolised. Particular attention should be paid to drugs used during organogenesis, i.e. between the third and eighth week after conception, that is from the fifth to tenth gestational week for most organs, and during the rapid fetal growth stage in the third trimester.

The monitoring system should include data on the drugs taken (generic name, dose, duration and exact time of administration in relation to gestation age). Data on the reproductive health outcomes of interest (diagnosis and treatments), is also needed, as well as an epidemiologic study design that allows computation of quantitative effect measures. The monitoring system should, therefore, be large and in most cases simple. Experience shows that the best achievable design depends as much on practical as on theoretical considerations.

2. Design Options

A monitoring system that aims at quantifying risks needs to obtain disease frequencies for exposed fetuses to contrast with disease frequencies for unexposed fetuses. In order to determine whether disease occurrences are related to drug expo-

sure, a follow-up study can be set up, or individuals with the disease of interest may be selected from the study base that gives rise to the cases, as it is done in a case-control study. Both designs aim to obtain the same relative effect measure from the same population.

Since most known drug-related adverse effects have a short incubation time and many alternative causes are stable over time and related to the mother, a monitoring system based on case-crossover methods could also be considered.

In principle, a monitoring system should include several potential consequences of drug use, and a follow-up study is the most obvious choice. However, if a special interest is taken in specific, serious and rare adverse effects, like congenital abnormalities, a case-control study may be the only possible choice.

It is not proposed that any of these monitoring systems replace existing reporting of spontaneous adverse drug reactions in medical practice. These reporting systems provide potentially important warning signals and state the need for risk assessment, which has to be based on epidemiological studies.

Most drugs are used worldwide and monitoring systems need not be set up everywhere. Areas with the best infrastructure for getting valid data should be selected. A time and place-specific monitoring system is only needed when adverse effects depend on local medical practice or genetic susceptibility present only in a specific population.

Follow-up may be done for the entire population of pregnant women or for selected exposed and unexposed women. Exposure data may be primary (generated for the project) or secondary (generated for other purposes). Follow-up usually starts in early pregnancy, but may start prior to conception if based upon routine drug registers that are operating all the time. Design options are many and only a few standard models are described in this paper, but the following discussion should be relevant to other design options as well. The most expensive design is discussed first and the low-cost models last.

3. Pregnancy Cohort, *Ad-Hoc* Studies

The model for such a study could be the Danish Pregnancy Cohort 'Better Health for Mothers and Children'.[9] Pregnant women are being recruited at their first visit to the general practitioner in gestational weeks 8 to 10. They are asked to respond to a short self-administered questionnaire on periconceptional drug use, and they are interviewed by phone at the end of the first and second trimester in order to give a complete history on the use of drugs during the first two thirds of the pregnancy. They provide two blood samples to a biological bank at the end of the first and second trimester and one from the umbilical cord shortly after delivery, which may be used as a source for biomarkers of drug use. Finally, 6 months after birth, mothers are questioned retrospectively about their drug use in the last trimester, and 18 months after birth they provide information on child development. The four interviews cover most of the known potential confounders for reproductive failures as well as the indications for drug use during pregnancy. Furthermore, a self-administered questionnaire on dietary intake is sent to the women around 24 weeks of gestation, including questions on vitamins and other supplements taken periconceptionally and during pregnancy. Data on adverse reproductive outcomes come mainly from routine registration systems and from the two postpartum interviews.

The study aims at including 100 000 pregnant women and will have more statistical power than the US perinatal collaborative study, which has been the largest follow-up study to date on drugs and pregnancy based upon primary data.[10] As at February 2001, 70 000 had been included in the Danish Pregnancy Cohort. A similar study has been initiated in Norway. However, a stand-alone cohort does not make a monitoring system in itself. Data collection has to be repeated at regular intervals, for example every 5 or 10 years, or to be an ongoing activity taking place in one or more countries.

Since the Danish Pregnancy Cohort includes personal identifiers, it is possible to use existing records as part of the monitoring system. Denmark has a routine registration of pregnancy and birth

complications, birthweight and length, head and abdominal circumference, and placenta weight. All hospital diagnoses are routinely computerised for mothers as well as for neonates, and the information is linkable to the cohort by means of these personal identifiers. The cohort study has data on potential confounders for congenital abnormalities, like diet and vitamins, as well as data on many other social factors that may play a role for other endpoints.

This procedure should ensure an unbiased exposure recording for the first two trimesters since outcomes data will not be known when medicine use is reported. There may be some nondifferential misclassifications of drugs, since drug names are difficult to recall, although information is supported by questions on the indications for treatment. Under-reporting of short-term treatment for minor diseases is also to be expected, but the only important likely source of differential recall bias is the previous reproductive history.[11] Early pregnancy complications may, however, have an influence on how drugs are recalled, and it may have impact on how affected women use these drugs.

This monitoring instrument provides possibilities for improving the diagnostic classifications of outcomes like congenital abnormalities by requesting medical records for those recorded with a congenital abnormality in the registry, and for those who reported a congenital abnormality at the interview. Subcohorts may be subject to a more detailed follow-up, for example cognitive performance in childhood. All children may be followed up in routine health registers for their entire life span. All of this is possible because mothers as well as offspring are identified by their personal identifiers.[12] Europe could use the unique large-scale registries of Nordic countries to obtain information that is probably not available in other countries.

This type of cohort is expensive and the estimated cost is about 10 million Euros in Denmark for data collection alone. The monitoring cost will, however, be much less when all data collection instruments have been developed and tested.

4. Pregnancy Cohort, Antenatal Care-Collected Data

Almost all countries in the developed world have a comprehensive antenatal care programme with screening schemes which often include too many visits, and which are rarely evidence-based.[13] Without jeopardising valuable clinical examination time, it should be possible to build up a computerised data file on drug intake. Such a monitoring system is in operation in Sweden, where midwives record drug use at the end of first trimester as part of the medical birth registration.[14] They routinely collect data on pregnancy complications and the birth outcome, and use the nationwide registry of congenital abnormalities.

The Swedish Medical Birth Registry was started in 1973 and contains computerised summaries of antenatal care records, delivery records, and records from the examination of the neonate (all neonates are examined by a paediatrician). Congenital malformations are identified from the Medical Birth Registry, the Registry of Congenital Malformations and the Child Cardiology Registry. At the first visit to the antenatal clinic (usually during weeks 10 to 12), the pregnant woman is interviewed by a midwife for about 1 hour,[15] where she is asked about drug use during the first trimester. Since 1995, data on all drugs taken after conception has also been collected and several analyses on specific drugs have now been published.[14,16-18]

Any misclassification of drugs taken is likely to be non-differential, although some drugs, especially over-the-counter drugs, may be under-reported. Nonresponses to the drug questions are unlikely to cause bias since data are recorded prospectively and therefore blinded to the outcome of the pregnancy. Practically all Swedish women attend these free antenatal clinics and approximately half a million neonates have been included in the Medical Birth Registry since 1995, with an addition of 80 000 to 100 000 neonates per year.

Information on potential confounders is limited, but personal identifiers make it possible to request medical records. By selecting a case-control study within this cohort, more detailed outcome meas-

ures could be made available. Adding these data to a routine antenatal care computerised database can be done at moderate cost, but the system depends upon good cooperation from midwives and general practitioners, which may be difficult to obtain. As in all studies using administrative databases and relying on discharge diagnoses, the completeness and validity of the information on drug use (exposure) and discharge diagnosis (outcome) must be evaluated.[19,20]

5. Pregnancy Cohort, Routine Pharmacy Registration of Exposure

Some countries run prescription databases, or dispensing databases, such as the computerised pharmacy data in The Netherlands. When these databases are linkable with routine birth registries,[21,22] subcohorts defined according to prescribed drugs may be identified at low cost. Frequency of adverse reproductive outcomes in these cohorts may then be compared with similar frequencies in women who have not been prescribed drugs. As in the antenatal care cohorts, nonresponders are not a problem since all pregnancies are registered.

The main problem is related to the proxy measure of drug intake. A prescribed drug, even if it is redeemed, need not be taken, or it may be taken by someone else. For some drugs, especially drugs taken for chronic diseases like asthma or migraine, the date of redeeming may be far removed from the date of use, which could be a serious problem when studying adverse effects that have a short time window of causal relevance. Furthermore, drugs which need no prescription are not included. Women classified as 'unexposed' according to the register may have taken over-the-counter drugs or even a prescription drug if it was prescribed before the pregnancy, or available from others.

In our experience, the prescription database works well for drugs prescribed for serious diseases that need continuous treatment. Since all drugs are registered for administrative reasons prior to the pregnancy outcome, any misclassification will be random, which usually will reducd risk estimates. Furthermore, drugs and their dosages are often registered accurately since this task is done by the prescribing physician or pharmacist. On the other hand, the possibility for confounder control is limited, and requesting additional medical records to check diagnoses may be prohibited.

Prescription databases have been used widely in pharmacoepidemiology. Still, relatively little is known about how well they describe drug intake, and the extensive literature on compliance is time and place-specific and is not necessarily generalisable to pregnant women. A monitoring system based upon prescription data should therefore be validated at regular intervals, and severe misclassification problems can be expected.[23]

The cost of such a monitoring system is low since data already exist. The cost of adding outcomes data may also be relatively inexpensive if data are computerised and identified by unambiguous identifiers like the personal identifiers.

Prescriptions are usually written by general practitioners and computerised data from this source have also been widely used in pharmacoepidemiology.[24,25] Since the prescriptions are not redeemed at the time of registration, another element of noncompliance adds to the uncertainty of the timing of drug use.

6. Case-Control Monitoring

A follow-up study on rare outcomes needs to be very large to produce sufficient information since most of the cohort members provide little information to the study. Only a fraction of cohort members is expected to have used a specific drug if they are not selected according to drug use, and most of the cohort members will not develop the disease under study. In our opinion, the only, but important, justification for using a cohort study is to get valid exposure data.

The prevalence of all congenital abnormalities is between 1 to 6% depending upon definitions and diagnostic routines, but specific types of congenital abnormalities do not exceed 5 to 15 per 1000 individuals for even the most frequent abnormalities such as congenital dislocation of the hip, ven-

tricular septal defect, neural tube defects, cleft lip/palate or clubfoot.[26,27]

A case-control study that tries to secure the necessary information by including all neonates with congenital abnormalities as well as a random sample of all neonates is usually much more efficient than a cohort study, measured by information units of cost. Furthermore, a proper sampling strategy allows unbiased estimates of relative effect measures without relying upon any rare disease assumption. A case-control sampling strategy requires, however, that the exposure in cases is reasonably prevalent, unless the exposure is a necessary cause for most of the newborns with the malformation under study.[3] In this situation a case-control study of rare exposure becomes informative. This will, however, be the exception rather than the rule.

In a case-control study of a congenital abnormality with a primary definition of the study base, usually a certain region or entire country, all neonates with the defined congenital abnormality should be ascertained to the case group. Controls should be obtained via a random sample of all neonates within the region and in the time period of study, but neonates without congenital abnormalities have often been used successfully since the disease is rare. The ratio of the exposure odds among cases divided with the exposure odds among controls provides an estimate of the relative prevalence ratio for congenital abnormalities; that is the prevalence proportion of congenital abnormalities among neonates exposed to the drug under study, divided by the prevalence proportion for congenital abnormalities among those not exposed. This relative 'cohort type' effect measure is reached by reconstructing the necessary exposure ratios in the sampling and it is slightly biased in a case non-case type of study, which is still the most common case-control study.

The Hungarian case-control monitoring system is one of the largest and best known monitoring programmes of congenital abnormalities in the world of the case–noncase type.[28-30] It is based upon national ascertainment of cases with a congenital abnormality and a sample of noncases. Between 1980 and 1996, 22 865 cases and 38 151 controls were recruited to the study, and the study is still ongoing although in a modified form.

The case-control monitoring permits control of the diagnostic procedures, and reclassification of cases according to all available medical documents could be implemented as part of the case ascertainment. This is especially necessary if one expects a more detailed search for congenital abnormalities among those exposed than among those not exposed. If physicians expect the treatment or its underlying disease (for example diabetes mellitus) to carry a higher risk of congenital abnormalities, closer medical surveillance may increase the risk of false-positive congenital abnormalities among exposed and perhaps also of false-negative congenital abnormalities among those not exposed, which will bias effect measures towards inflated values.

The Achilles heel of case-control monitoring is the reconstruction of exposures for the correct time period, which is usually several months before case ascertainment and interviewing. If this recording is based upon recall, it usually means remembering events at least 6 to 10 months back in time. Some exposures will not be recalled, and, more seriously, incentive to recall will produce differential recall in some situations.[31,32] Evidence indicates better recall for cases than for controls in most situations,[33] but the recall may be opposite if the drugs are illegal, not prescribed or not considered necessary.

In some case-control studies, it may be possible to get access to exposure data recorded prospectively during pregnancy. In the Hungarian case-control monitoring system, medical records (log books) filled in by physicians treating pregnancy-related complications are also collected. Log book data cover only a fraction of all treatments, but probably provide an unbiased exposure since the data are recorded prior to the diagnosis.

A case-control study of this type is based upon informed consent and nonresponding may produce selection bias. At present, it is difficult to run a case-control study with more than 70 to 80% participation, which is sufficiently low to cause serious bias

in a sensitive field like use of medicine, and in many countries the participation rates may even be lower. In the Hungarian case-control monitoring system, 80% of cases provided information on drug use compared with 65% of controls, since the ethics committee only permitted a 'reminding' visit to the former group. This decision has repercussions for the validity of the inference made.

Ongoing monitoring systems of congenital abnormalities, such as the MADRE (Malformations and Drug Exposures) and EUROCAT (European Registration of Congenital Anomalies) systems, may be used for surrogate case-control monitoring if data on drug intake are included in the monitoring.[34-36] If exposure data are collected successfully, specific congenital abnormalities could then serve as cases and other congenital abnormalities as patient controls, if it is reasonable to assume that these controls have diseases that are neither caused by nor prevented by the drug in question. Experience from the Hungarian database, where neonates with Down's syndrome were used as an alternative control group, indicates that symmetry in recall is partly obtained by using patient controls. Since data are collected for the project, it is possible to add information on potential confounders to the case-control analysis.

Even in large case-control studies, like the Hungarian study, statistical power is limited for drugs that are not used frequently. Only few studies are large enough to provide meaningful results for specific congenital abnormalities. Some grouping of congenital abnormalities is often needed and the options are described in section 8.

7. Classification of Congenital Abnormalities

Congenital anomalies span from inborn errors of metabolism to functional disorders and include growth retardation. Most registries only include congenital abnormalities, i.e. the consequence of morphogenesis errors defined as a structural defect, severe or mild, present at birth whether it is detected at that time or not. Previously, an anatomical-based classification was applied, which probably only partially correlated with a classification based upon aetiologic mechanisms.

A compromise between the anatomical and the aetiological classification could be to use a pathogenetic approach by separating congenital abnormalities into isolated or nonsyndromic and multiple syndromic congenital abnormalities. The isolated congenital abnormalities are morphologic defects, which can be traced back to one localised error of morphogenesis, while multiple congenital abnormalities are the result of two or more different morphogenetic errors. However, the identification of isolated or multiple congenital abnormalities in a given congenital abnormality entity needs rather detailed information and expert evaluation. It is an important task to differentiate congenital abnormalities according to their severity, also since severity is expected to correlate with completeness of ascertainment.

The following groupings are often used: (i) lethal – prenatally detected fetal defects followed by termination of pregnancy; (ii) severe – without medical intervention the congenital abnormality is life-threatening and/or leads to severe handicap (i and ii are called major congenital abnormalities); and (iii) mild – congenital abnormality that may need medical treatment but without it, is not life threatening, for example undescended testis and dislocation of the hip.

Congenital abnormalities may also be classified according to the gestational time of origin. Furthermore, a classification could be based upon the origin of cells that produce the congenital abnormality, for example crista neuralis–derived organs have been grouped together because these cells have a more rapid cell division, migrate over longer distances and therefore may be more vulnerable to external exposures.[37]

Most studies on congenital abnormalities estimate the proportion of abnormalities at the time of birth, which is a frequency at a given point in time, and is thus a prevalence measure. The limitation of this measure is that it only records abnormalities of those who survive until birth and as a consequence, teratogens that always lead to abortions will

never be detected by using this measure. In addition, there is a growing proportion of prenatally detected fetal defects followed by pregnancy termination, which will cause bias if this is done because of the drug exposure. Another problem is that the prenatal diagnostics may be done with a test that has false-positives and false-negatives. If the sensitivity and specificity of the test are known, the monitoring may incorporate proper correction factors into the estimates if induced abortions based upon prenatal diagnostics are included.[38,39]

8. Case-Time-Control Monitoring

Many methodological problems would be solved if cases could be used as their own controls, especially if the response rate among controls is a problem. Confounding by time-stable factors is eliminated and, in some cases, so is confounding by an underlying disease. Confounding by indication is still a problem if the reproductive failure under study is caused by fluctuation of the disease rather than treatment. This is a problem in all designs without randomisation of drug use.

In its simplest form, the monitoring may be based upon a case-crossover design. Cases may, for example, be preterm births, and exposures closely related in time to this event are compared for a similar exposure window – 1 week, 1 month, or perhaps for a previous pregnancy. If one wants to study whether use of a given asthma treatment activates labour prematurely, intake of the drug may be recorded in a time period of 48 hours prior to the onset of labour and compared with a similar 48-hour time period the week before. By adding up cases who were exposed prior to labour but not in the reference period, and by dividing this figure with the number of cases not being exposed prior to labour but in the reference period, an estimate is obtained of how much more frequently women go into labour shortly after the drug use, although interpretation is saddled with restriction. First of all, the possibility that the onset of labour activated the asthma attack needs to be ruled out. If it cannot be ruled out, results are confounded by indication, as labour is the cause of treatment.

The use of drugs should be independent of the duration of pregnancy, and be stable over the time period of study. If not, a proper reference group is needed for adjustment.[40,41] Furthermore, participation in the study must be independent of the pattern of drug use. If, for example, women who used the drug prior to labour are more likely to join the study than women who used the drug in the reference period, selection bias will occur.

Case-crossover principles may also be used when studying congenital abnormalities. Exposure during the organogenesis may be compared with exposure outside this time window. If we study limb defects for which the relevant time period of organogenesis is the eighth gestational week, a reference period in gestational week 12 may be selected. A reference period before week 8 is more bias-prone since the pregnancy may not yet have been detected. On the other hand, it is also possible for multiparous women to compare drug use in the pregnancy that produced the case with drug use in a previous pregnancy. If the treatment was given for a chronic disease, which was present in both pregnancies, and if the use of the drugs changed over the time period under study, a case-time control adjustment should be used which calls for sampling of controls.

As for a case-control monitoring system, it is restricted to the outcomes used in selecting cases, but it permits adjustment for confounders and quality control of the diagnosing is possible. The case-crossover design should be considered a possible alternative to case-control monitoring if nonresponding controls is a problem.

9. Discussion and Recommendations

A monitoring system of drug use during pregnancy should ideally be able to identify short-term and long-term consequences of drug use and be able to provide independent estimates of adverse effects from different sources. For this reason several designs and data sources are needed. Analysis and interpretation of these data should be coordinated, which is not the case at present.

When choosing a monitoring system, feasibility as well as validity has to be considered. Validity has first priority, and ideally one would want a system sensitive enough to pick up all adverse effects, important either because of their severity or their frequency. These findings would then be scrutinised using a system design with high specificity and as few false positive results as possible, although some will be unavoidable. No single system will provide this, and it is necessary to combine several monitoring systems to achieve both a high sensitivity and a high specificity. A follow-up monitoring system is expected to be specific but with low sensitivity. For a case-control study with self-repeated data the opposite can be expected, a high sensitivity but a low specificity. A monitoring system should be seen as a set of evaluations based on different data sources and design options.

A prescription database may provide accurate data for treatments of unexpected acute diseases or chronic disorders that require ongoing drug therapy. It may be the monitoring system of choice in these situations, but cannot stand alone. Many drugs are not associated with high rates of compliance and prescription databases provide very few options to control for confounding. Since most reproductive failures have a short time window of exposure, a registration system that does not get the timing right will be biased towards negative findings, which means a low sensitivity. Accurate exposure assessment will require repeated questioning, which will be expensive unless healthcare personnel can record this data. Pregnant women are in close and frequent contact with the healthcare system, and it should be possible to implement a cohort with antenatal care-collected data as an ongoing monitoring system in many countries. Use of memory aids and self-administered questionnaires filled out by pregnant women at home would greatly enhance credibility of such a system.

Few monitoring systems address spontaneous abortions as an outcome. Many abortions occur in the very early stages of pregnancy and it is difficult to recruit women to a cohort study before conception. Studies on spontaneous abortion may therefore be best addressed in a monitoring system based on prescription cohorts if computerised hospital records have data on induced abortions. Since induced abortions may distort the outcome measure, the effect of induced abortions should be controlled, if possible.[38]

Even very large follow-up studies will still fall short in detecting rare outcomes. Well-functioning case-control monitoring systems should, therefore, not be abolished but rather encouraged to continue and to improve. The time of recall of drug use should be equal for cases and controls and should be as close to the time of birth as possible. A case-control study should focus upon serious diseases like congenital abnormalities, but could possibly also include childhood cancers, cerebral palsy, autism and mental retardation. A case-control monitoring system should also allow case-crossover analyses, which should perhaps replace case-control analyses in many situations. We consider the case-crossover study to be a very promising new development in pharmacovigilance in this field.

We also believe that study of large pregnancy cohorts should be included in monitoring systems. This should be established in a country that is able to run such a study with long and complete follow-up. Only this study will be able to investigate functional disorders that are not routinely recorded,[42] and provide the possibility of using biomarkers of exposure and susceptibility.

Since all these data collection systems are in operation, the task ahead is to strengthen coordination while using existing experience. We have used our experience from EuroMaP in writing this review. However, there are several other monitoring systems in existence,[43-49] which could be incorporated into the system.

There are good data to show that fetotoxicity may have more than one manifestation, which speaks strongly in favour of the cohort type of monitoring. In particular, data are needed on childhood behaviour and cognitive functions, which will be very difficult to collect by means of case-control sampling. Long-term follow-up must be included in a modern

monitoring programme on drug safety for pregnant women and their unborn children.

All of the available designs have shortcomings. Therefore, positive as well as negative associations should be evaluated in another data set, preferably with a different design. We recommend setting up a coordinated system in different countries since there is a large variation in availability of drugs on the market, prescription practices vary and genetic traits of importance are not uniformly distributed.[19] It should be possible to identify regions best suited to run a specific monitoring system to the benefit of all. We suggest that all monitoring systems produce databases with free access for external researchers.

Such a monitoring system needs approval from high-level health authorities. There is no point in running a monitoring system unless there is a willingness to use the results; in other words, to turn the monitoring system into a surveillance system.

Questions on adverse effects of drugs often refer to drugs used during pregnancy.[50] Unfortunately, it is not clear who has the responsibility to provide the information and pay the expenses for getting the best evidence-based answers. Ongoing systems to provide the best available data on exposure and outcomes are needed as a basis for counselling women. New drugs are coming on the market that sooner or later will be used by pregnant women and even information with limited statistical power on which to draw conclusions is valuable.[51]

Principles of data protection have often been used to justify the lack of monitoring for drug-related adverse effects, and protecting data seems often to be more important than protecting people. Prior experience shows that monitoring can be done without unwanted disclosure of personal data. After appropriate linkage, monitoring can be done on data without personal identifiers since results are based upon comparisons between groups and not individuals.

Even large monitoring systems collect sparse information. At present this information is often published in minor journals if results are negative, or results are not published at all and readers of literature get a highly biased view of the evidence.[52,53] We suggest that all results are sent to an international database whether positive or negative, and whether they are published or not. The organisation responsible for the registration system should then publish the combined results at regular intervals.

10. Conclusion

In conclusion, experience is available to set up a comprehensive monitoring system of drug-related adverse effects for offspring exposed to drugs *in utero*, and it is time to do so. A package of different designs is proposed that cover an in-depth cohort study with long-term follow-up and a focus on drug-induced organ programming. We also suggest the inclusion of a prescription database study to screen for possible associations that need closer scrutiny in other studies. Furthermore, the case-time-control approach to search for drug-related severe adverse effects should be evaluated. All of these research activities need to be coordinated.

Acknowledgements

This study was supported by the EuroMaP concerted action in the Biomed 2 work programme, contract no. BMH4 CT97-2430. Corinne S. de Vries has been funded by a Wellcome Trust Travelling Research Fellowship, grant no. 557509/Z/99/Z. The activities of the Danish Epidemiology Science Centre are financed by a grant from the Danish National Research Foundation.

References

1. Bonati M, Bortolus R, Marchetti F, et al. Drug use in pregnancy: an overview of epidemiological (drug utilization) studies. Eur J Clin Pharmacol 1990; 38: 325-8
2. Marchetti F, Romero M, Bonati M, et al. Use of psychotropic drugs during pregnancy: a report of the international cooperative drug use in pregnancy (DUP) study. Collaborative Group on Drug Use in Pregnancy (CGDUP). Eur J Clin Pharmacol 1993; 45: 495-501
3. Lenz W, Knapp K. Thalidomide embryopathy. Arch Environ Health. 1962; 5: 100-5
4. McBride WG. Thalidomide and congenital abnormalities [letter]. Lancet 1961; ii: 1358
5. Koren G, Pastuszak A. Drugs in pregnancy. N Engl J Med. 1998; 338: 1128-37
6. Rubin P. Drug treatment during pregnancy. BMJ 1998; 317: 1503-6

7. Olsen J. Prenatal exposures and long term health effects. Epidemiol Rev 2000; 22: 76-8
8. Giusti RM, Iwamoto K, Hatch EE. Diethylstilbestrol revisited: a review of the long-term health effects. Ann Intern Med 1995; 122 (10): 778-88
9. Olsen J, Melbye M, Olsen SF, et al. The Danish national birth cohort: its background, structure and aim. Scand J Public Health 2001; 29: 300-7
10. Heinonen OP, Slone D, Shapiro S. The women, their offspring, and the malformations. In: Birth defects and drugs in pregnancy. Littleton (MA): Publishing Sciences Group, 1977: 30-3
11. Olsen J. Options in making use of pregnancy history in planning and analysing studies of reproductive failure. J Epidemiol Community Health 1994; 48: 171-4
12. Frank L. When an entire country is a cohort. Science 2000; 287: 2398-9
13. Carroli G, Villar J, Piaggio G, et al. WHO systematic review of randomised controlled trials of routine antenatal care. Lancet 2001; 357 (9268): 1565-70
14. Källén B. Delivery outcome after the use of acid-suppressing drugs in early pregnancy with special reference to omeprazole. Br J Obstet Gynaecol 1998; 105: 877-81
15. Cnattingius S, Ericson A, Gunnarskog J, et al. A quality study of a medical birth registry. Scand J Soc Med 1990; 18: 143-8
16. Kallen B, Lygner PE. Delivery outcome in women who used drugs for migraine during pregnancy with special reference to sumatriptan. Headache 2001; 41 (4): 351-6
17. Källén B, Rydhström H, Åberg A. Congenital malformations after the use of inhaled budesonide in early pregnancy. Obstet Gynecol 1999; 93: 392-5
18. Ericson A, Källén B, Wiholm B-E. Delivery outcome after the use of antidepressants in early pregnancy. Eur J Clin Pharmacol 1999; 55: 503-8
19. Bergman U. Pharmacoepidemiological perspectives. J Clin Epidemiol 1992; 45: 313-7
20. Calle EE, Khoury MJ. Completeness of the discharge diagnoses as a measure of birth defects recorded in the hospital birth record. Am J Epidemiol 1991; 134: 69-77
21. Olesen C, Steffensen FH, Sørensen HT, et al. Pregnancy outcome following prescription for Sumatriptan. Headache 2000; 40: 20-4
22. Steffensen FH, Nielsen GL, Sørensen HT, et al. Pregnancy outcome with ACE-inhibitor use in early pregnancy [letter]. Lancet 1998; 351: 596
23. de Vries CS, de Walle HEK, Cornel MC, et al. Registration of drug use in a birth defect monitoring system: a priority worthy of emphasis! Int J Risk Saf 1993; 4: 27-34
24. Lau HS, deBoer A, Beuning KS, et al. Validation of pharmacy records in drug exposure assessment. J Clin Epidemiol 1997; 50: 619-25
25. Jick H, Holmes LB, Hunter JR, et al. First-trimester drug use and congenital disorders. JAMA 1981; 246: 343-6
26. Stoler JM. Prevalence and etiology of birth defects. In: Prenatal care. McCormick MC, editor. Cambridge: Cambridge University Press, 1999
27. Czeizel AE, Intödy ZS, Modell B. What proportion of congenital abnormalities can be prevented? BMJ 1993; 306: 499-503
28. Czeizel AE. Epidemiological studies of congenital abnormalities in Hungary. In: Kalter H, editor. Issues and reviews in teratology. Dordrecht: Kluwer Academic Publishers, 2000: 85-124
29. Czeizel AE, Ráez J. Evaluation of drug intake during pregnancy in the Hungarian case-control surveillance of congenital anomalies. Teratology 1990; 42: 505-12
30. Czeizel AE. The first 25 years of the Hungarian congenital abnormality registry. Teratology 1997; 55: 299-305
31. Mitchell AA, Rosenberg L, Shapiro S, et al. Birth defects related to Bendectin use in pregnancy. JAMA 1981; 245: 2311-4
32. Bar-Oz B, Moretti ME, Mareels G, et al. Reporting bias in retrospective ascertainment of drug-induced embryopathy. Lancet 1999; 354: 1700-1
33. Rockenbauer M, Olsen J, Czeizel AE, et al. Recall bias in a case-control surveillance system on the use of medicine during pregnancy. Epidemiology 2001; 12: 461-6
34. de Jong-van den berg LTW, Feenstra N, Sørensen HT, et al. Improvement of drug exposure data in a registration of congenital anomalies. Teratology 1999; 60: 33-6
35. de Vries CS, de Walle HEK, Cornel MC, et al. Registration of drug use in a birth defect monitoring system: a priority worthy of emphasis! Int J Risk Saf Med 1993; 4: 27-34
36. de Jong-van den berg LTW, Waardenburg CM, Haaijer-Ruskamp FM, et al. Drug use in pregnancy: a comparative appraisal of data collecting methods. Eur J Clin Pharmacol 1993; 45: 9-14
37. Hansen D, Lou HC, Olsen J. Serious life events and congenital malformations: a national study with a complete follow-up. Lancet 2000; 356: 875-80
38. Olsen J. Calculating risk ratios for spontaneous abortions: the problem of induced abortions. Int J Epidemiol 1984; 13: 347-9
39. Garne E, Bergman U. Benzodiazepine use in pregnancy and major malformations or oral clefts: induced abortions should be included [letter]. BMJ 1999; 319: 918
40. Greenland S. Confounding and exposure trends in case-crossover and case-time-control designs. Epidemiology 1996; 7: 231-9
41. Suissa S. The case-time-control design: further assumptions and conditions. Epidemiology 1998; 9: 441-5
42. Reinisch JM, Sanders SA, Mortensen EL, et al. In utero exposure to phenobarbital and intelligence deficits in adult men. JAMA 1995; 274: 1518-25
43. Grisso JA, Carson JL, Feldman HI, et al. Epidemiological pitfalls using Medicaid data in reproductive health research. J Matern Fetal Med 1997; 6 (4): 230-6
44. Rosa F. Databases in the assessment of the effects of drugs during pregnancy. J Allergy Clin Immunol 1999; 103 (2 Pt 2): S360-1
45. Jick SS. Pregnancy outcomes after maternal exposure to fluconazole. Pharmacotherapy; 19 (2): 221-2
46. Bar Oz B, Hackman R, Einarson T, et al. Pregnancy outcome after cyclosporine therapy during pregnancy: a meta-analysis. Transplantation 2001; 71 (8): 1051-5
47. Vargas FR, Schuler-Faccini L, Brunoni D, et al. Prenatal exposure to misoprostol and vascular disruption defects: a case-control study. Am J Med Genet 2000; 95 (4): 302-6
48. Hernandez-Diaz S, Werler MM, Walker AM, et al. Folic acid antagonists during pregnancy and the risk of birth defects. N Engl J Med 2000; 343 (22): 1608-14
49. Ruigomez A, Garcia Rodriguez LA, Cattaruzzi C, et al. Use of cimetidine, omeprazole, and ranitidine in pregnant women

and pregnancy outcomes. Am J Epidemiol 1999; 150 (5): 476-81
50. Öhman B, Lyrvall H, Törnqvist E, et al. Clinical pharmacology and the provision of drug information. Eur J Clin Pharmacol 1992; 42: 563-8
51. Källén B. Drugs in pregnancy: the dilemma of labeling. Drug Inf J 1999; 33: 1135-43
52. Brent RL. Computer program for searching the biomedical literature in teratology, developmental biology, and reproductive biology. Teratology 1982; 25 (3): 273-5
53. Mastroiacovo P, Serafini M, Pagano M, et al. The 'Telefono Rosso': a service for the prevention of birth defects and for the evaluation of teratogenic risk. Ann Ist Super Sanita 1993; 29 (1): 115-20

Correspondence and offprints: Dr *Jørn Olsen*, The Danish Epidemiology Science Centre at the Department of Epidemiology and Social Medicine, University of Aarhus, Vennelyst Boulevard 6, Aarhus, 8000 Aarhus C, Denmark.
E-mail: jo@soci.au.dk