REVIEW ARTICLE

Drug Safety 2008; 31 (1): 39-51
0114-5916/08/0001-0039/$48.00/0

Data Resources for Investigating Drug Exposure during Pregnancy and Associated Outcomes

The General Practice Research Database (GPRD) as an Alternative to Pregnancy Registries

Rachel A. Charlton,[1] *Marianne C. Cunnington*,[1] *Corinne S. de Vries*[2] and *John G. Weil*[1]

1 Worldwide Epidemiology, GlaxoSmithKline, Harlow, Essex, UK
2 Department of Pharmacoepidemiology and Pharmaceutical Medicine, Postgraduate Medical School, University of Surrey, Guildford, UK

Contents

Abstract

Pregnancy registries are the most commonly used data resource for the post-marketing surveillance of drug teratogenicity. However, the limited sample size and potential selection bias in these registries has led us to investigate the potential of the UK General Practice Research Database (GPRD) as an alternative data source for monitoring drug safety during pregnancy. In addition, a literature review identified further observational data sources that monitor pregnancy outcomes for future evaluation.

Initial feasibility studies focused on the ability of the GPRD to capture pregnancy outcomes for a range of drug class exposures, all of which are currently under investigation in pregnancy registries, during pregnancy. The comparator pregnancy registries were identified via a MEDLINE search, whilst eligible pregnancies, in which women received one or more prescriptions for the drug of interest during pregnancy, were identified in the GPRD using the mother-baby

link. The number of pregnancy outcomes following exposure to medication for a range of conditions with varying prevalence, including depression, migraine, epilepsy, herpes simplex and HIV, captured by the two data sources were compared. For depression, a relatively prevalent condition, the GPRD recorded the same number of mean annual intrauterine exposures to fluoxetine as the pregnancy registry (118 exposures/year). Ascertainment of intrauterine exposure to drug treatments for less prevalent conditions was found to be higher for the pregnancy registries than the GPRD; for the older antiepileptic drugs (valproate and carbamazepine), the pregnancy registry recorded between four and five times as many mean annual exposures as the GPRD. Virtually no antiretroviral exposures (three) were identified during the time period of interest on the GPRD, compared with 3946 in the Antiretroviral Pregnancy Registry.

Data from the GPRD meet established criteria for evaluating outcomes of pregnancy. For prevalent conditions, it has the potential to replace or work alongside standard pregnancy registries and the alternative data sources identified. Further studies are now needed to assess its ability to replicate known teratogenic associations.

Prescription medications are commonly used by women of childbearing age.[1] Since it is estimated that almost half of pregnancies in the US are unplanned,[2] there is the possibility of the fetus being exposed to medication during the critical period of organ and tissue development known as morphogenesis.[2] Even if a pregnancy is planned, certain medical conditions such as epilepsy or depressive disorders can make it inadvisable for women to discontinue their medications given the risk that the underlying condition poses to the health of both mother and child.

Pregnant women are rarely included in clinical trial programmes for ethical reasons; hence, data on the potential teratogenicity (i.e. the ability to cause malformations in the developing fetus) or other developmental effects on the fetus of new medical treatments are often limited to those obtained from animal experiments.[3] However, the risk in animals does not always accurately predict the risk in humans,[4] which means that the safe use of a drug during pregnancy cannot be totally assessed until the drug is on the market, making post-marketing surveillance a vital source of information.

At present, pregnancy registries are the most common form of post-marketing surveillance for the detection of drug teratogenicity, and are commonly seen as an essential part of an ongoing programme in epidemiological safety monitoring. Registries are voluntary, national or international prospective studies. They aim to detect signals of major teratogenicity, such as in the case of thalidomide (in which 25% of exposed infants are malformed), through the detection of substantial increases in the occurrence of all malformation types. Registries can monitor single drugs or a drug class and are usually developed by the pharmaceutical industry sponsor or by university-based research groups. Prospective enrolment (before the pregnancy outcome is known) is a key methodological strength of pregnancy registries as retrospective reporting (after the pregnancy outcome is known) can bias towards the capture of more severe outcomes.

GlaxoSmithKline currently sponsors four prospective pregnancy registries, the longest of which has been running since 1989, yet no GlaxoSmithKline registry has managed to reach the enrolment milestone of 1000 pregnancies, considered by the European Committee for Medical Products for Human Use (CHMP) to be representative of widespread exposure, during the first 10 years of marketing of the medication. In addition to low enrolment, selection bias can be introduced through the voluntary nature of enrolment and loss to follow-up. This

potential bias, combined with limitations in statistical power resulting from small sample sizes, restrict the power of a registry to the identification of large increases in risk. This led us to investigate the potential of the UK General Practice Research Database (GPRD) as an alternative data source to pregnancy registries. Numerical comparisons focused on the ability of the GPRD to capture outcomes for a range of drug-class exposures during pregnancy. An additional literature review was completed to identify any further data sources using a cohort approach capable of being used for the monitoring of outcomes of pregnancy.

1. Data Sources and Methodology

The GPRD is the world's largest computerized database of anonymous, primary-care medical records.[5] The database currently contains information on approximately 3 million active patients from some 400 general practices within the UK. Virtually all prescriptions, non-drug interventions and referrals issued by general practitioners (GPs) are recorded in the database, as are medical diagnoses, including pregnancy. The GPRD contains a mother-baby link (still to be officially validated) that enables the medical records of mothers to be linked to the medical records of their offspring if their child is born alive and is registered with the same practice. This is based on a practice-specific family identification number (primarily based on address), confirmed by ensuring that the mother's delivery date is within ± 60 days of the child's date of birth.

Our analysis of alternative data sources for the monitoring of drug teratogenicity, in terms of the ascertainment of intrauterine drug exposure, focuses on a comparison of the GPRD with drug- or indication-specific pregnancy registries. Data sources were identified through a MEDLINE search using the key search terms 'pregnancy registry', 'pregnancy outcome', 'abnormalities', 'abnormalities, drug-induced' and 'teratogens'. The search was limited to dates between January 1990 and March 2006 (inclusive) and to drugs licensed for use in the UK. Registries were required to report details of first-trimester exposure, a detailed methodology describing how individuals were classified as drug-exposed and the time period of data collection. The latest interim reports of the GlaxoSmithKline-sponsored pregnancy registries were also accessed.

All mother-baby pairs identified through the mother-baby link on the GPRD, with a birth outcome between January 1991 and October 2005 and where the mother was aged between 14 and 49 years at the inferred delivery date, were evaluated with respect to medications under surveillance in the registries identified above. As the mother-baby link had been introduced only 2 years earlier, January 1991 was selected as the start date for GPRD data extraction due to the very low number of mother-baby pairs identified prior to this. Those pairs in which the mother had not been followed for the entire pregnancy period, taken as the 280 days prior to the date of the birth, were excluded, as were pairs registered at practices where the data provided was not considered by the Medicines and Healthcare products Regulatory Agency (MHRA) to be of sufficient standard for research purposes.

In the GPRD, exposure to a drug during pregnancy was defined as the mother receiving one or more prescriptions for the drug of interest during the time period of 90–280 days prior to the date of the birth. This time period was selected as a best estimate in view of the difficulties in precisely identifying the pregnancy start and end dates resulting from the frequent unavailability of the date of the last menstrual period for pregnant women in the GPRD, and the range in gestational weeks that a pregnancy can last. Both trade names and generic drug names were searched to ensure that data were captured for as many of the exposed mother-baby pairs as possible. The definitions of 'exposure' used in the individual pregnancy registries, in terms of monotherapy and polytherapy within a drug class and any additional inclusion criteria, were used to enable a direct comparison of exposure ascertainment to be made between the pregnancy registries and the GPRD.

The time period for which data was extracted from the GPRD often differed from the time period for which the pregnancy registries reported expo-

sure data. To account for this, the comparison of exposure ascertainment was based on the mean number of pregnancy exposures recorded per year. This was calculated by dividing the total number of exposures identified by the number of years of data collection. Where drug treatments were first marketed after the start date of data extraction from the registry or GPRD, the calculations for the mean number of annual exposures were based on the month and year the drug was marketed for use.

2. Comparison of the GPRD with Pregnancy Registries

Within the GPRD, 377 408 mother-baby pairs with a live birth outcome between January 1991 and October 2005 were identified. Of these, 229 788 had been followed for the entire pregnancy period, were permanently registered with the practice, and the quality of data for both the mother and baby was considered to be of sufficient standard by the MHRA.

Pregnancy registry data were available on drug treatments for a range of conditions with varying prevalence (epilepsy,[6,7] HIV,[8] migraine,[9] depression[10] and herpes virus infection[11]). All registries included in the analysis, with the exception of the UK Epilepsy and Pregnancy Registry, were international and included data from a number of countries. In some cases, however, the reporting to international registries was dominated by one particular country or region, for example, 89% of reports to the International Antiretroviral Pregnancy Register were from the US and 3.8% of reports from the UK.[8] With the exception of registries for anticonvulsants, only one pregnancy registry per condition was available for comparison with the GPRD. Comparisons with the UK Epilepsy and Pregnancy Register[6] were of particular interest as the registry has a similar population base to the GPRD.

A comparison of the number of exposed pregnancy outcomes captured by pregnancy registries and the GPRD for the above conditions and medications is described in table I. For the drug fluoxetine, used to treat depression, the pregnancy registry[10] and the GPRD data sources captured the same number of mean annual intrauterine exposures. The number of mean annual intrauterine exposures to acyclovir for the treatment of herpes virus infection was greater in the GPRD than the international pregnancy registry[11] (121 vs 40 exposures/year).

Overall, exposure ascertainment of treatments for epilepsy, HIV and migraine was lower in the GPRD than for the pregnancy registries and, proportionally, newer anticonvulsants had a higher representation in the UK Epilepsy and Pregnancy Register[6] than in the GPRD (where representation mirrored the market share). In the case of lamotrigine, this could be in part due to the fact that the GPRD data collection started immediately after the drug's launch, before it had had time to penetrate the market. Further analysis showed that no intrauterine exposures to lamotrigine are recorded on the GPRD prior to 1996. Comparison of the mean number of annual exposures for the same time period as the UK Epilepsy and Pregnancy Register still found exposure ascertainment to be considerably lower in the GPRD than the pregnancy registry (8.3 and 77.6 exposures/year, respectively). For the older antiepileptic drugs, such as valproate and carbamazepine, the UK-based epilepsy register[6] recorded four to five times as many mean annual intrauterine exposures as the GPRD. In the case of antiretroviral drugs, a total of three intrauterine exposures to any monotherapy or polytherapy antiretroviral regimen were identified in the GPRD compared with 3946 reported to the international Antiretroviral Pregnancy Registry.[8] Based on the size of the US population, the fact that 3.8% of registry reports came from the UK and the fact that the GPRD covers approximately a 5% sample of the UK population, this is less than half the number of exposures to antiretroviral drugs than would be expected.

3. Other Available Data Sources

Table II summarizes the additional data sources, identified via the MEDLINE search, that captured exposure to medications during pregnancy and corresponding outcomes. Data sources that used a cohort approach largely collected exposure data prospectively and were grouped as follows: observa-

Table I. Ascertainment of intrauterine drug exposure on the General Practice Research Database (GPRD) compared with pregnancy registries

Drug exposure and indication	Date range	No. of exposures in:		Mean no. of exposures/year in:	
		pregnancy registry	GPRD (Jan 1991–Oct 2005)	pregnancy registry[a]	GPRD[a]
Epilepsy[b]		UK Epilepsy and Pregnancy Register[6]			
Valproate	1 Dec 1996–31 Mar 2005	715	270	85.8	18.2
Lamotrigine	1 Dec 1996–31 Mar 2005	647	82	77.6	5.9
Carbamazepine	1 Dec 1996–31 Mar 2005	900	431	108.0	29.1
Gabapentin	1 Dec 1996–31 Mar 2005	31	13	3.7	1.0
Levetiracetam	1 Nov 2000–31 Mar 2005	22	0	5.1	0.0
Topiramate	1 Dec 1996–31 Mar 2005	28	4	3.4	0.4
Phenytoin	1 Dec 1996–31 Mar 2005	82	62	9.8	4.2
Migraine		Naratriptan/Sumatriptan Pregnancy Registry[9][c]			
Naratriptan	1 Oct 1997–30 Apr 2006	38	78	4.4	9.2
Sumatriptan	1 Jan 1996–30 Apr 2006	372	184	36.0	13.6
Depression		Eli Lilly and Company Worldwide Fluoxetine Pregnancy Registry[10]			
Fluoxetine	1 Jul 1989–9 Apr 1999	796	757	117.9	118.4
Antiretroviral		International Antiretroviral Pregnancy Registry[8][c,d]			
Abacavir	1 Jan 1999–31 Jan 2006	345	0	48.7	0.0
Lamivudine	1 Nov 1995–31 Jan 2006	1663	1	163.6	<0.5
Zidovudine	1 Jan 1989–31 Jan 2006	1459	1	85.4	<0.5
Nevirapine	1 Jun 1996–31 Jan 2006	479	1	56.4	<0.5
Herpes virus infection		Acyclovir and Valacyclovir Pregnancy Registry[11]			
Acyclovir	1 Jun 1984–30 Apr 1999	597	1798	40.2	121.2
Valacyclovir	1 Jan 1995–30 Apr 1999	22	8	5.1	0.7

a For those where the drug became marketed after the start date of data collection, the mean numbers of exposures per year are based on the month/years the drug was first marketed.

b In monotherapy.

c Sponsored by GlaxoSmithKline and contracted out to an independent research organization.

d Previously known as the Zidovudine in Pregnancy Registry. Became the International Antiretroviral Pregnancy Registry in 1993.

Table II. Description of data sources available to assess the effect of drug exposure during pregnancy

Data source	Time period of data collection	Population coverage	Additional risk information (all capture maternal demographics)	Source of exposure and outcome information
Observational studies with voluntary recruitment (excluding drug and/or indication specific pregnancy registries)				
Organisation of Teratology Information Services[12]	Since 1985	Consists of 45 teratology information services in US and Canada	Medication use, medical histories may be reviewed in some cases, some variation by service	Largely self-reporting of exposure and outcome via a maternal telephone interview although some variation by service
European Network of Teratology Information Services[13]	Since 1990	Consists of teratology information services in >14 European countries and Argentina and Brazil	Varies by service	Self-reporting of exposure data. Outcome information requested from either the mother or her physician via a mailed questionnaire or telephone call
Population-based surveillance registries				
Swedish Medical Birth Register[14]	Since 1994	~98% of Swedish population; 83 300–117 600 births/year	Social health (smoking, alcohol), medical data (including BMI and BP) and medication use	Routine self-reporting of exposure data via maternal interview with midwife Outcome recorded by paediatrician and obtained via hospital records or congenital malformation register Personal identifiers make it possible to request medical records
Norwegian Medical Birth Registry[15]	Since 1967; drug exposure data since 1998	Compulsory reporting covers all births; ~60 000 births/year	Occupation, smoking, folate intake, medication use, maternal medical history	Exposure data recorded on an antenatal form during visits to GP or midwife and brought to the hospital by the mother Outcomes recorded by midwives and physicians
Finnish Medical Birth Register[16]	Since 1987	Compulsory reporting covers all births; ~58 000 births/year	Occupation, maternal medical use	Exposure data collected via antenatal care visits Outcome data recorded by hospital personnel

Continued next page

Table II. Contd

Data source	Time period of data collection	Population coverage	Additional risk information (all capture maternal demographics)	Source of exposure and outcome information
Denmark National Registry of Patients (in parallel with nationwide registers of malformations)[17]	Reporting of malformations since 1963, but this method since 1995	Compulsory reporting covers all births; ~50 000 births/year	Occupation, maternal medical use	Routine self-reporting of exposure data via maternal interview Outcome recorded by paediatrician and routinely computerized Opportunity to request original medical records
Saskatchewan population registries[18]	Since 1970	93% of Canadian province of Saskatchewan population; ~11 400 births/year	Some social status indicators, medication use	Drug plan provides data on all prescriptions written to the Saskatchewan population Outcome/hospital data are recorded by the physician electronically Primary health records can be accessed either by visiting physician clinics or physician self-reporting
European Concerted Action on Congenital Anomalies and Twins (EUROCAT)[19]	Since 1980	Network of 43 congenital anomaly registers in 20 countries. ~25% of births in Europe; ~1.2 million births/year	Varies by registry Drug exposure data is not collected by all registries	Ability to access and the method of access to exposure data varies by register Outcome largely reported by physician to local/national congenital anomaly register Access to original medical records varies by register
Healthcare databases				
General Practice Research Database[20]	Since 1986	Currently, 5% sample of 60 million UK population and those who have died/been transferred out; ~59 000 births/year	Medical and treatment history	Exposure and outcome data recorded in medical records by GP Access to free text comments, original medical records and ability to send questionnaires to GPs are available at a charge
Kaiser Permanente[21]	Since ~1995	San Francisco area, US; ~30 000 births/year	Medication use	Exposure and outcome data via medical and pharmacy claims Access to original medical records available at a charge

Continued next page

Table II. Contd

Data source	Time period of data collection	Population coverage	Additional risk information (all capture maternal demographics)	Source of exposure and outcome information
Tennessee Medicaid[22]	Since 1985	~36 000 births/year	Pregnancy medication history	Exposure and outcome data via medical and pharmacy claims Access to original medical records available at a charge
United Healthcare[23]	Since 1990	US; ~27 000 births in 1997	Medical and pharmacy claims	Exposure and outcome data via medical and pharmacy claims Access to original medical records available at a charge
Case-control studies				
National Birth Defects Prevention Study[24]	Since 1997	Currently 10% of annual US birth cohort; 7470 cases and 3821 controls in 2000	Social, medical history, medication and environmental exposures during pregnancy	Self-reporting of exposure via maternal telephone interview Outcome recorded in medical records by paediatrician and reviewed by clinical geneticist
Slone Epidemiology Unit Birth Defects Study[25]	Since 1976 (depends on drug of interest)	Covers births in the US/Canadian metropolitan areas of Boston, Toronto, Philadelphia, San Diego; ~12 000 cases and 4000 controls in 2003	Social, medical history, medication and environmental exposures during pregnancy	Self-reporting of exposure via maternal interview Outcome recorded by paediatrician Access to original medical records with mother's permission
Hungarian Case Control Surveillance of Congenital Abnormalities Study[26]	Since 1980	Covers all births in Hungary; 22 843 cases and 38 151 controls in 1996	Medical history, medication exposures during pregnancy	Obstetrician records exposure in antenatal logbook and additional information requested in a retrospective questionnaire Physician or obstetrician records the outcome Access to antenatal logbook and discharge summaries are available
Estudio Colaborativo Latino Americano de Malformaciones Congential[27]	Since 1967	Latin America; ~150–200 000 births/year	Medical history, drug intake and other prenatal exposures	Exposure and outcome information collected from the mother by a trained paediatrician during puerperium

BMI = body mass index; **BP** = blood pressure; **GP** = general practitioner.

tional studies with voluntary recruitment; population-based surveillance registers; and healthcare databases.

All cohort studies based on voluntary recruitment were based within teratology information services such as the Organization of Teratology Information Services (OTIS) and European Network Teratology Information Services (ENTIS) that offer counselling to individuals concerned about aspects of their pregnancy. These use telephone help lines as a method for recruiting cohorts of women who have been exposed to particular drugs during pregnancy.

The population-based surveillance registers can be further divided into comprehensive or linked registers, comprehensive registers providing a single surveillance system for the reporting of details of exposure to medication during pregnancy and reporting of birth outcomes (e.g. the Swedish Birth Medical Register). Reporting is often mandatory in countries with these registers. Linked registers require an additional step involving unique patient identifiers (often as part of health insurance programmes) to link information in separate registers monitoring maternal drug exposure (e.g. pharmacy registers) and the health of the infant after birth (e.g. congenital malformation registers) at the national level. Information on the formal validation of mother-baby linkage was not usually described for these surveillance systems.

In addition, several medical claims insurance databases from the US collecting pregnancy exposure and outcome data were identified.

4. Observations

The starting point for this investigation was the recognition of recruitment limitations and potential selection bias inherent in the pregnancy registry design. We have evaluated the potential ascertainment of drug exposure in the GPRD and listed complementary data sources that can be used to evaluate teratogenicity (table II).

The comparison of drug exposure ascertainment in pregnancy registries and in the GPRD makes it apparent that the level of recording of intrauterine drug exposure in the GPRD is dependent on disease prevalence, the length of time a drug has been on the market, and UK prescribing practice. For example, depression and migraine are two relatively prevalent conditions in the UK (affecting >3% and >10% of women of child-bearing age, respectively), yet while the GPRD proved a potentially useful additional source of data to monitor teratogenic potential of antidepressants, very few exposures to triptans for acute treatment of migraine were identified. This may reflect UK prescribing practice which rarely treats triptans as a first-line migraine therapy, but is in agreement with work carried out in 2004 by Bakker et al.[28] Their study observed a decrease in the rate of prescriptions for antimigraine medication during the 3 months prior to pregnancy and during the pregnancy period, with an increase during the 3 months post birth. They hypothesized that this could be due to women experiencing considerably fewer attacks during the period of pregnancy or the fact that women may take alternative analgesics such as paracetamol (acetaminophen), thought to be safe to use during pregnancy. Although the mother-baby link within the GPRD excludes those pregnancies that do not result in live births (e.g. resulted in abortions), this is unlikely to have significantly impacted upon the estimated total number of exposures in pregnancy to a particular drug treatment.

For drug treatments prescribed for conditions with a lower prevalence (e.g. HIV and epilepsy), pregnancy registries are clearly still the more efficient and informative approach, providing information on a considerably greater number of mean annual exposures than the GPRD. This may reflect the fact that treatment for those conditions is organized through tertiary referral units, resulting in underrecording of exposures in the GPRD, in addition to the larger populations from which the registries are recruiting.

Further data sources with the potential of being used for monitoring pregnancy outcomes following drug exposure were also identified and evaluated. Previous studies have discussed additional data sources for monitoring drug safety during pregnancy,[29] but in this review we focus specifically in terms of seven main criteria emphasized in US FDA

guidance.[30] These criteria are discussed in more details in sections 4.1–4.7 below.

4.1 Representativeness

Capturing a representative sample is important, especially where the population is heterogeneous in terms of factors that might modify the underlying association with disease outcome, for example, disease severity or ethnicity in relation to specific defects. A failure to capture such high-risk groups or their over-representation could lead to erroneous conclusions when applied at the population level.

The GPRD and national registries have an advantage over pregnancy registries in terms of capturing the general population or a representative sample of the population. In the case of pregnancy registries, it is harder to access representative samples of the population because they rely on voluntary participation. However, within the GPRD, only 60% of pregnancies met the inclusion criteria and were followed for the entire pregnancy period. In addition, because of the restrictions in the actual definition of the mother-baby link (e.g. only live births are included), a large proportion of all pregnancies potentially of interest have not been identified. Consequently, there is the potential for selection bias that needs to be investigated further and pregnancies not resulting in live births need to be considered for inclusion.

4.2 Capture of Exposure Data

Self-reported exposure, as in some pregnancy registries, national surveillance registers (e.g. those in Sweden) and cohorts from teratology information services, raises the question of information accuracy, especially in terms of the dosage and duration of exposure to a particular drug. The Swedish register, relying on self-reported exposure data, has estimated that only 60–70% of antiepileptic drug use was reported.[14] Such under-reporting may introduce selection bias, which may be a problem when investigating early, first-trimester events and under-reporting is also likely to have statistical power implications, especially in the case of rarer treatments.

The GPRD has the advantage that exposure data are recorded prospectively and routinely by the prescriber. However, even when exposure data are reported through the physician, an assumption needs to be made that the medication prescribed was actually taken and taken close to the time the prescription was written. For chronic conditions like epilepsy, this seems a fair assumption to make; however, for episodic conditions, such as migraine, where drugs are frequently prescribed in anticipation of a future acute event, the dates of prescribing may not necessarily mark the beginning of actual drug use.[31] In addition, a Danish study found that only 43% of drugs (of any kind) dispensed to pregnant women were reported to be taken[32] and therefore there is the issue of misclassification of exposure. In the pregnancy registries, this is more likely to be differential, leading to an overestimation of risk, whereas in databases it will be non-differential and likely to lead to an underestimation of risk.

4.3 Loss to Follow-Up of Women and Infants

The voluntary nature of recruitment to pregnancy registries can result in loss to follow-up, which has the potential to introduce bias if cases lost to follow-up were to differ systematically in terms of their malformation risk to those remaining in the register. In the GPRD, loss to follow-up is only possible due to patients moving practices or practices ceasing to contribute to the database. National surveillance registers can sometimes reduce loss to follow-up through mandatory reporting and linkage of multiple outcome registers as in the case of birth, hospital and congenital malformation registers in Sweden for example.

4.4 Ascertainment of Outcome Data

The completeness and accuracy of the infant medical report depends on the reporter and the level of detail may vary widely. Pregnancy registries are likely to provide the most accurate information because most cases are reviewed by an expert in teratology and the reporter is usually the obstetrician attending the birth or the infant's paediatrician. The GPRD, as well as other population-based registers, are more limited in this regard and further work is needed to validate specific outcomes of interest

within the database. One advantage of the GPRD, however, is the opportunity to request and obtain additional information recorded in the free text comments section of the infant's medical record and to send questionnaires to the patients' GPs, which will allow such validation to be carried out.

The reporting of malformations at birth restricts studies to malformations that have survived to birth, those that have not been detected by prenatal screening and those that are detected at birth. Advances in prenatal screening means that diagnoses of malformations are often made earlier in pregnancy which, for some individuals, results in the decision to have a pregnancy termination. The extent to which data sources capture these cases varies and it is therefore essential that when reporting rates of congenital malformations that both the numerator and denominator are clearly defined. The current mother-baby link in the GPRD only covers live births and further work is needed to evaluate how comprehensively data can be collected on induced abortions in the database. However, this issue is common to all data sources considered, and the lack of systematic evaluation of all induced abortions and stillbirths remains an important limitation. One further limitation in terms of ascertainment of outcome is related to the variable proportion of fatal pregnancy outcomes for which autopsies are carried out. This, again, is an issue common to all the data sources considered.

Finally, the adverse effects of drugs during pregnancy are not necessarily limited to teratogenic effects. While both the pregnancy registries and the purpose-built monitoring systems tend to focus on congenital anomalies because they cover all routinely recorded medical history, such as birth weight and gestational age, healthcare databases offer the opportunity to evaluate other effects such as intrauterine growth and/or early onset of labour.

4.5 Data on Additional Factors and Influences on Pregnancy Outcome (Confounders)

Information on maternal risk factors varies greatly between sources. The effects of very high-risk teratogens are likely to outweigh concerns surrounding adjustment for confounders, and can still be informative. Additional data on potential confounders do, however, prove beneficial when quantifying actual risk associated with moderate increases in risk. Pregnancy registries often encourage recruitment and follow-up by minimizing the effort for the physician and participant in terms of paperwork and time, which may mean the collection of less additional information. This trend commonly extends to national surveillance registers, although in Sweden antenatal health interviews by midwives collect extensive data on covariates. The GPRD contains data on additional drug exposures during pregnancy as well as maternal health. The level of detail and accuracy of reporting (e.g. smoking status, alcohol intake) may vary by covariate.

4.6 Internal Control Group

For single-drug pregnancy registries, there is no internal control group, although one can argue that an appropriate control group is not possible (as women with a given condition not on medication are likely to differ inherently from those on medication). The problem can be limited by comparisons with a general population group and cohorts from the literature of women exposed to other monotherapies.

Healthcare databases such as the GPRD enable access to internal control groups, which should be subject to the same recruitment biases as the exposed group of interest. Multiple internal controls are possible, including women exposed to other drugs in the same class, women with the same condition not taking medication and general population controls. National surveillance systems tend to rely on the general population for comparison; however, if the underlying disease itself increases the risk of congenital malformations then the general population will not be an ideal control and lead to overestimated risks.

4.7 Statistical Power

The statistical power to detect an existing increase in risk is related to the frequency of outcome

and/or exposure (depending on study design) as well as the size of the relative risk to be detected. We have already discussed the issues surrounding sample size and factors that might influence the sample size captured, but in terms of relative risk, pregnancy registries are best suited for the detection of large increases in risk or the detection of increases in all malformation types. This commonly limits pregnancy registries to the generation of hypotheses related to teratogenicity associated with increases in the risk of specific defect types representing rare outcomes (e.g. a baseline risk of 1 in 1000). The ability of the GPRD to further the aims of pregnancy registries appears to be highly dependent on the condition being studied. The case-control surveillance design, however, has greater power to detect moderate teratogenicity. It is for this reason that it is recommended that cohort studies are initiated to exclude major teratogenicity followed by the case-control approach, used to complement and test hypotheses around specific malformations.[33]

5. Conclusion

One of the main aims of teratogen surveillance is to collect information and make it available to pregnant women and GPs as soon as possible after a drug is first launched. Pregnancy registries remain the most frequently used surveillance method; when they are set up internationally, they are the most likely to collect information on large numbers of pregnancies relatively quickly. However, as drugs remain on the market for longer periods of time, reporting to such registries decreases and for prevalent conditions of a chronic nature the GPRD has the potential to replace or work alongside standard pregnancy registries. For less prevalent conditions, the potential remains for the GPRD to continue surveillance and replicate findings once the pregnancy registry has acquired the cohort of 200 exposed newborns required to provide some assurance that the drug in question is not a major teratogen.[34] However, further validation of the mother-baby link, refinement in determining the timing of exposure, information on specific birth outcomes and the potential to capture outcomes for induced abortions are needed prior to studies demonstrating the GPRD's ability to replicate known teratogenic associations. Alternative data sources, including national surveillance registers, also exist and may not be known to all researchers and are therefore underutilized.

Acknowledgements

The authors thank Dr Janet Cragan for her contribution and comments to the earlier drafts of this paper and also Deepak Sriramulu for his assistance with the analyses.

No sources of funding were used in the preparation of this review article. Miss Charlton and Drs Cunnington and Weil are employees of and hold shares in GlaxoSmithKline. Dr de Vries has no conflicts of interest that are directly relevant to the content of this review article.

References

1. Honein MA, Paulozzi LJ, Cragan JD, et al. Evaluation of selected characteristics of pregnancy drug registries. Teratology 1999; 60 (6): 356-64
2. Henshaw SK. Unintended pregnancy in the United States. Fam Plann Perspect 1998; 30 (1): 24-9
3. Koren G, Pastuszak A, Ito S, et al. Drugs in pregnancy. N Engl J Med 1998; 338 (16): 1128-37
4. Hoar R. Developmental toxicity: extrapolation across species. J Am Coll Toxicol 1995; 14: 11-20
5. Hardy JR, Holford TR, Hall GC, et al. Strategies for identifying pregnancies in the automated medical records of the General Practice Research Database. Pharmacoepidemiol Drug Saf 2004; 13 (11): 749-59
6. Morrow J, Russell A, Guthrie E, et al. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. J Neurol Neurosurg Psychiatry 2006; 77 (2): 193-8
7. Kendle International Inc. Lamotrigine pregnancy registry [online]. Available from URL: http://www.kendle.com/registries [Accessed 2007 Nov 22]
8. Antiretroviral Pregnancy Registry Steering Committee. Antiretroviral Pregnancy Registry International Interim Report for January 1989 through 31 July 2005 [online]. Available from URL: http://www.kendle.com/registries [Accessed 2007 Nov 22]
9. Kendle International Inc. Sumatriptan and Naratriptan Pregnancy Registry International Interim Report. 1 January 1996 through 31 October 2005 [online]. Available from URL: http://www.kendle.com/registries [Accessed 2007 Nov 22]
10. Goldstein DJ, Corbin LA, Sundell KL. Effects of first-trimester fluoxetine exposure on the newborn. Obstet Gynecol 1997; 89 (5): 713-8
11. Glaxo Welcome. Acyclovir Pregnancy Registry and Valacyclovir Pregnancy Registry. Final Study Report. 1 June 1984–30 April 1999 [online]. Available from URL: http://pregnancyregistry.gsk.com/acyclovir.html [Accessed 2007 Nov 22]
12. Leen-Mitchell M, Martinez L, Gallegos S, et al. Mini-review: history of organized teratology information services in North America. Teratology 2000; 61 (4): 314-7

13. Schaefer C, Amoura-Elefant E, Vial T, et al. Pregnancy outcome after prenatal quinolone exposure. Evaluation of a case registry of the European Network of Teratology Information Services (ENTIS). Eur J Obstet Gynecol Reprod Biol 1996; 69 (2): 83-9
14. Centre for Epidemiology. The Swedish Medical Birth Register: a summary of content and quality, 2003 [online]. Available from URL: http://www.sos.se/fulltext/112/2003-112-3/2003-112-3.pdf [Accessed 2006 Jun 30]
15. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. Acta Obstet Gynecol Scand 2000; 79 (6): 435-9
16. Hemminki E, Teperi J, Tuominen K, et al. Need for and influence of feedback from the Finnish birth register to data providers. Qual Assur Health Care 1992; 4 (2): 133-9
17. Christensen K, Knudsen LB. Registration of congenital malformations in Denmark. Dan Med Bull 1998; 45 (1): 91-4
18. Government of Saskatchewan. Saskatchewan Health. Saskatchewan's health services information resources, July 2002 [online]. Available from URL: http://www.health.gov.sk.ca/programs-services [Accessed 2007 Nov 22]
19. Dolk H. EUROCAT: 25 years of European surveillance of congenital anomalies. Arch Dis Child Fetal Neonatal Ed 2005; 90 (5): F355-8
20. Garcia Rodriguez L, Perez-Gutthann S, Jick S. The UK General Research Practice Database. In: Strom B, editor. Pharmacoepidemiology. Chichester: John Wiley, 2000: 375-85
21. Li DK, Liu L, Odouli R, et al. Exposure to non-steroidal anti-inflammatory drugs during pregnancy and risk of miscarriage: population based cohort study. BMJ 2003; 327 (7411): 368
22. Hartert TV, Neuzil KM, Shintani AK, et al. Maternal morbidity and perinatal outcomes among pregnant women with respiratory hospitalizations during influenza season. Am J Obstet Gynecol 2003; 189 (6): 1705-12
23. Cole JA, Modell JG, Haight BR, et al. Bupropion in pregnancy and the prevalence of congenital malformations. Pharmacoepidemiol Drug Saf 2007; 16 (5): 474-84
24. Yoon PW, Rasmussen SA, Lynberg MC, et al. The National Birth Defects Prevention Study. Public Health Rep 2001; 116 Suppl. 1: 32-40
25. 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
26. Czeizel AE, Petik D, Vargha P. Validation studies of drug exposures in pregnant women. Pharmacoepidemiol Drug Saf 2003; 12 (5): 409-16
27. Castilla EE, Orioli IM, Orioli IM. ECLAMC: the Latin-American collaborative study of congenital malformations. Community Genet 2004; 7 (2-3): 76-94
28. Bakker MK, Jentink J, Vroom F, et al. Drug prescription patterns before, during and after pregnancy for chronic, occasional and pregnancy-related drugs in The Netherlands. BJOG 2006; 113 (5): 559-68
29. Olsen J, Czeizel A, Toft Sørensen H, et al. How do we best detect toxic effects of drugs taken during pregnancy? Drug Saf 2002; 25 (1): 21-32
30. US Department of Health and Human Sciences Food and Drug Administration, Center for Drug Evaluation and Research, and Centre for biologics Evaluation and Research. Reviewer guidance. Evaluation of human pregnancy outcome data, 1999 [online]. Available from URL: http://www.fda.gov/cder/guidance/6777fnl.htm#_TOC79910680 [2006 Sep 30]
31. Velentgas P, Cole JA, Mo J, et al. Severe vascular events in migraine patients. Headache 2004; 44 (7): 642-51
32. Olsen C, Sondergaard C, Thrane N, et al. Do pregnant women report use of dispensed medications? Epidemiology 2001; 12: 497-501
33. Mitchell AA. Systematic identification of drugs that cause birth defects: a new opportunity. N Engl J Med 2003; 349 (26): 2556-9
34. Covington DL, Tilson H, Elder J, et al. Assessing teratogenicity of antiretroviral drugs: monitoring and analysis plan of the antiretroviral pregnancy registry. Pharmacoepidemiol Drug Saf 2004; 13 (8): 537-45

Correspondence: Dr *Marianne C. Cunnington*, Worldwide Epidemiology, GlaxoSmithKline, Third Avenue, Harlow, Essex, CM19 5AW, UK.
E-mail: marianne.c.cunnington@gsk.com