Methodological challenges in using routinely collected health data to investigate long-term effects of medication use during pregnancy

Luke E. Grzeskowiak, Andrew L. Gilbert and Janna L. Morrison

Abstract: To date, the investigation of teratogenic effects of medications has largely focused on physical alterations present at birth (i.e. malformations) as opposed to functional alterations (i.e. neurodevelopment, metabolic function) that may not be apparent at birth but could influence an individual's health and risk of disease in later life. The use of routinely collected health data represents one approach to better identifying, quantifying, and understanding the long-term risks or benefits of medication use during pregnancy. As such, the objective of this review was to identify and explore opportunities and challenges associated with using routinely collected health data to examine long-term effects of medication use during pregnancy. Drawing on published research several key methodological issues associated with their use in investigating long-term outcomes are reviewed. While significant opportunities exist to make greater use of routinely collected health data, there are a number of key challenges. Identified challenges relate to aspects of study design and analysis, and include obtaining access to data, the ability to match records across datasets and over long periods of time, how medication exposures are ascertained and classified, issues around loss to follow-up how outcomes are ascertained and classified, and the careful interpretation of results in light of study and data limitations. Understanding key challenges associated with using routinely collected health data to investigate long-term effects of medication use during pregnancy is essential in supporting their appropriate use and interpretation, which will contribute to improving the quality of research undertaken and ensure the reliability of results obtained.

Keywords: data linkage, medication, pharmacovigilance, pregnancy, teratogen

Introduction

The frequent exclusion of pregnant women from randomized controlled trials places a heavy reliance on observational studies to gather evidence on any associated risks and benefits of medication use during pregnancy. To date, the investigation of teratogenic effects of medications has largely focused on physical alterations (i.e. malformations) as opposed to functional alterations (i.e. neurodevelopment, metabolic function) that may not be apparent at birth but could influence an individual's health and risk of disease in later life.

A conceptual framework for understanding the contribution that maternal medication use during

pregnancy could have on the health and development of the offspring in later life can be taken from the rapidly emerging and expanding field of the developmental origins of health and disease [McMillen and Robinson, 2005]. Through the combination of experimental, clinical, epidemiological, and public-health research, there is an aim to understand how events in early life influence an individual's susceptibility towards health or disease in later life [McMillen and Robinson, 2005]. While much of the focus of this research is on early life exposures to factors other than medications (i.e. maternal diet), little is known about the potential long-term effects of medication use during pregnancy. The discovery of an

Ther Adv Drug Saf

(2013) 4(1) 27–37

DOI: 10.1177/

2042098612470389

© The Author(s), 2012.

Reprints and permissions:
[http://www.sagepub.co.uk/
journalsPermissions.nav](http://www.sagepub.co.uk/journalsPermissions.nav)

Correspondence to:

Luke E Grzeskowiak,
BPharm (Hons),
GCertClinEpi

Quality Use of Medicines
and Pharmacy Research
Centre, School of
Pharmacy and Medical
Sciences, Sansom Institute
for Health Research,
University of South
Australia, GPO Box 2471,
Adelaide, South Australia
5001, Australia

Luke.Grzeskowiak@
mymail.unisa.edu.au

Andrew L. Gilbert,
BPharm, DipAppPsych,
PhD

Professor, Quality Use of
Medicines and Pharmacy
Research Centre, School
of Pharmacy and Medical
Sciences, Sansom Institute
for Health Research,
University of South
Australia, Adelaide, South
Australia, Australia

Janna L. Morrison, BSc
(Hons), MSc, PhD

Heart Foundation South
Australian Cardiovascular
Network Fellow, Early
Origins of Adult Health
Research Group, School
of Pharmacy and Medical
Sciences, Sansom Institute
for Health Research,
University of South
Australia, Adelaide, South
Australia, Australia

association between the use of diethylstilbestrol during pregnancy and an increased rate of vaginal adenocarcinomas in young women exposed *in utero* is a pertinent reminder of the need for adequate follow-up of all medication exposures during pregnancy [Herbst *et al.* 1971]. More recently, potential long-term effects following prenatal exposure to other medications have emerged. For example, there is evidence that *in utero* exposure to sodium valproate is associated with an increased risk of impaired cognitive function in the child at 3 years of age [Meador *et al.* 2009], and that prenatal antidepressant exposure is associated with an increased risk of autism spectrum disorders (ASD) [Croen *et al.* 2011], and disturbed development of the enteric nervous system [Nijenhuis *et al.* 2012], in childhood.

While it is recognized that no single approach is sufficient for delineating the complete spectrum of risks or benefits of medication use in pregnancy [Mitchell, 2003], the use of routinely collected health data represents one approach to better identifying, quantifying, and understanding the long-term risks or benefits of medication use during pregnancy. The major advantage of this approach is that data are already available, collected prospectively in a routine fashion (e.g. as prescription records or hospital discharge summaries), stored electronically, and can be linked at the individual patient level [Holman *et al.* 1999]. In addition, use of routinely collected health data enables the identification of very large cohorts of women and their offspring, which is particularly useful in assessing infrequent, but clinically important, outcomes [Grzeskowiak *et al.* 2012a]. Linkage of existing records is unobtrusive, and the routine and prospective manner in which data are collected and recorded in these datasets eliminates recall bias [Holman *et al.* 1999]. In addition, as data have already been collected, studies can be undertaken in a more timely and cost-efficient manner than is possible with other research design methods. While there are numerous potential benefits in using routinely collected health data, methodological challenges associated with its use in examining long-term effects of medication use during pregnancy have not been well described. As such, the objective of this review was to identify and explore opportunities and challenges associated with using routinely collected health data to examine long-term effects of medication use during pregnancy. Drawing on published research using routinely collected health data we review several key validity and

methodological issues associated with their use in investigating long-term outcomes. While significant opportunities exist to make greater use of routinely collected health data, there are a number of key challenges. In particular, key methodological challenges relate to aspects of study design and analysis, and include obtaining access to data, the ability to match records across datasets and over long periods of time, how medication exposures are ascertained and classified, issues around loss to follow-up how outcomes are ascertained and classified, and the careful interpretation of results in light of study and data limitations. A summary of strengths and limitations of using routinely collected health data to examine long-term effects of medication use during pregnancy are outlined in Table 1, with associated challenges discussed in more detail below.

Access to data

Data security and confidentiality are two important issues associated with the use of routinely collected health data. Importantly, these issues can be addressed and appropriate protocols and procedures have been implemented in research centers, within Australia and internationally, to ensure this type of research can be undertaken in an ethical manner [Kelman *et al.* 2002]. As such, while ethics approvals and/or data custodian approvals are required to obtain data, they do not prevent appropriately designed and managed projects from using routinely collected health data. Furthermore, a recent review of research outputs from the Western Australian data-linkage system highlighted not only the enormous amount of research that has been undertaken using linked data, but was also able to demonstrate the public health benefits of this research through resultant improvements in policy and/or practice [Brook *et al.* 2008]. At present there is immense potential and opportunity to investigate long-term outcomes associated with medication use during pregnancy, much of which is just starting to be realized. Subsequently, it can be argued that we have an ethical obligation to utilize data that has already been collected if it may result in a greater understanding of the long-term effects of medication use during pregnancy, and lead to improved health outcomes for mothers and their offspring.

Data linkage

A challenge in undertaking research using routinely collected health data is the ability to link

Table 1. Strengths and limitations of using routinely collected health data to examine long-term effects of medication use during pregnancy.

Strengths	Limitations
Can answer a variety of research questions in a relatively timely and efficient manner	Restricted to investigating exposure and outcomes routinely recorded in the data
Commonly population-based, improving generalizability of findings and reducing potential for selection bias	Potential misclassification of medication exposures and outcomes
Datasets often contain data on many individuals and can therefore generate large sample sizes	Assumptions regarding dose, timing, and duration of medication exposure based on dispensing data
Data routinely collected as part of clinical practice/administration requirements	Diagnostic data often not available (i.e. severity of underlying illness)
Unobtrusive data collection	Available data limited to that which is routinely collected. Relevant data of interest (i.e. potential confounders) not collected
Eliminates potential for recall bias	Data may differ across data sources (i.e. differences in the way data are collected in different settings or changes to data collection over time)
Enables long-term follow-up	Data quality and integrity may differ across data sources
Useful in assessing rare, long-term outcomes	Large sample size can lead to increased likelihood of chance findings

data pertaining to the same individual, in this case mother–child pair, over long periods of time. The data-linkage process can be undertaken using either deterministic or probabilistic matching. Deterministic matching is used when records can be directly linked using a unique identifier, which may take the form of a national identification number (i.e. CPR in Denmark), or a health insurance ID number. The inclusion of a unique identifier, such as the unique 10 or 11 digit code assigned to citizens in Nordic countries [Kieler, 2010], enables highly accurate and efficient linkage of records across population-based datasets.

In the absence of a unique identifier, linkage must occur using probabilistic matching, which involves linking records using a number of personal identifiers such as surname, first given name, date of birth, sex, and address [Kelman *et al.* 2002]. The difficulty encountered with longitudinal and cross-generational matching is accounting for changes in name (e.g. through marriage or divorce), or address, and any spelling/typographical errors. In addition, some people may be lost to follow-up as they change their state or country of residency. Despite these difficulties, very large population-based datasets have been generated using probabilistic matching within Australia [Kelman *et al.* 2002], Canada [Chamberlayne *et al.* 1998], the UK [Hardy *et al.*

2006], and the USA [Manning *et al.* 2011]. These have been used in many research projects to date. Importantly, while growing project experience and advances in computer programming make the linkage of records less of a technical barrier, there is still the need to consider potential factors that may be associated with incomplete data linkage, and whether these factors may introduce bias in reported clinical outcomes [Bohensky *et al.* 2010, 2011]. For example, linkage of Medicaid claims and vital records (i.e. birth-certificate data) in the USA has previously been demonstrated to underrepresent outcomes of high-risk pregnancies, potentially introducing selection bias [Bronstein *et al.* 2009]. An additional limitation of using claims databases from healthcare plans such as Medicaid, which is a healthcare plan for low-income persons, is that cohort members are not representative of those with private health insurance, limiting the generalizability of study findings [Cole *et al.* 2007b]. Similarly, claims-based databases from private health insurers suffer from the limitation that analyses are restricted to those who maintain coverage with the health insurer during the defined study period, which could change as a result of changes in income, employment, and/or residence. In addition, the generation of mother–child pairs is often dependent on the child being covered under the same insurer as the mother. In a US-based

study, only 75% of mothers and their children were covered under the same insurer [Cole *et al.* 2007a]. This is similar to the reported percentage of identified mother–child pairs in the German pharmacoepidemiological research database (77.3%), which is also based on insurance-claims data [Garbe *et al.* 2011].

Exposure ascertainment and classification

A number of key methodological issues exist when using electronic dispensing data to classify medication exposure during pregnancy. With significant differences often observed across different studies [Grzeskowiak *et al.*, 2012b]. A previous methodological review of studies on antidepressant use in pregnancy demonstrated that key factors such as the dose, duration, and timing of exposure were inconsistently addressed, and there was a great deal of variability in the way medication exposures were classified, and how women who stop taking their medication before, or during early pregnancy, are handled in analyses [Grzeskowiak *et al.*, 2012b]. Further issues identified related to studies not having data on gestational age to determine exact timing of exposure, and assumptions regarding how and when women who receive a dispensing for a medication actually take it during pregnancy. This creates a great deal of uncertainty around actual medication exposure during pregnancy, potentially resulting in biased risk estimates. These issues are discussed in more detail below.

Understanding medication-use patterns

Before using electronic dispensing data to classify medication exposures during pregnancy it is important to understand medication-use patterns. Using selective serotonin reuptake inhibitors (SSRIs) as an example, studies have demonstrated that up to 40% of women discontinue their SSRI before, or very soon after, finding out that they are pregnant [Bakker *et al.* 2008]. This is important to understand as it means that up to a third or more of women who receive a dispensing for an SSRI at or around the time of conception may not actually take that SSRI during the pregnancy. Therefore, studies that classify women as exposed if they receive a dispensing in the month prior to conception are potentially subject to misclassification bias, as some women could be misclassified as exposed when they were not actually taking the medication during pregnancy [Grzeskowiak *et al.* 2012b]. This bias has the potential to result in an

underestimation of the risk associated with medication use in pregnancy.

Timing of exposure

It is essential that studies investigating outcomes associated with medication use during pregnancy carefully consider and identify exposures occurring during relevant periods of fetal development that are likely to be associated with the outcome of interest. The majority of studies undertaken using routinely collected health data have focused on associations between prenatal medication exposure and congenital malformations. In these cases the critical periods of exposure have been well studied and defined during the first trimester. In contrast, for studies looking at long-term outcomes (i.e. effects on child neurodevelopment or metabolic function), the critical period of exposure is often not so well defined and may be much wider. Therefore, it is important to have a sound understanding of expected biologically relevant periods of exposure. Databases that include estimated dates of last menstrual period, and/or gestational age of the infant at delivery, which is confirmed by ultrasound, have the significant advantage of being able to evaluate outcomes according to timing of exposure. However, there are additional challenges in determining the association between when a medication is dispensed and when it is taken.

Assumptions regarding medication use

The underlying assumption associated with all studies relying on prescription-claims databases is that a dispensing for a medication is actually associated with use of that medication. Noncompliance with medication use may result in misclassification bias, where some of the women classified as ‘exposed’ were in fact not exposed during pregnancy. For example, Olesen and colleagues compared reporting of medication use during pregnancy among women enrolled in the Danish National Birth Cohort to the North Jutland prescription database, for those women living in the North Jutland area [Olesen *et al.* 2001]. Medications used for chronic medical conditions (i.e. diabetes, epilepsy, hypothyroidism) were always reported to be used by women (100%), whereas sensitivity was poor for medications used for acute conditions (e.g. 40% for nonsteroidal anti-inflammatory drugs). Overall, sensitivity for use of any medication was 43% (95% confidence interval [CI] 40–46). The additional assumption made when using dispensing data is that the date of dispensing reflects the time at which the medication was taken. This may be an appropriate

assumption for medications used for chronic medical conditions, but may not be valid for medications used on an intermittent basis for acute medical conditions.

It is important to note that as long as data on medication exposures during pregnancy are collected without bias, incomplete capture of medication exposure is unlikely to have a significant effect on the final risk estimate [Grzeskowiak *et al.* 2012b; Kallen, 2005]. A more significant issue is the need to ensure those who are considered exposed were actually taking the medication during pregnancy, as this can introduce significant bias and alter the final risk estimate [Grzeskowiak *et al.* 2012b]. There is a clear need for further research to determine the most effective and accurate way of using routinely collected health data to investigate outcomes following medication use during pregnancy in an effort to minimize potential biases.

Caution with dose

A limitation associated with many electronic dispensing datasets is that they only give an indication of the strength of the medication dispensed, but not the dose or frequency of use. This can make it difficult to investigate the effect of dose on various outcomes and the dose-response relationship. Previously, researchers have estimated dose by dividing the amount of medication supplied by the number of days between each supply [Berard *et al.* 2007]. However, while evidence of a dose-response relationship strengthens arguments for establishing causality, relying on estimates of dose may not provide the complete picture. For example, key factors such as individual differences in pharmacokinetics (i.e. due to differences in body weight, concurrent medication use, placental transfer) and pharmacogenomics play important roles in influencing the concentration and effects of fetal exposure to particular medications. As such, the same dose administered to two different individuals may elicit different responses. In many cases these types of data are not available in electronic records, and can only be examined further through additional data collection.

Following a sound understanding of the importance of the way in which medication exposures are classified during pregnancy, it is possible to move towards the consideration of various electronic datasets that could be used to identify medication exposures.

Medication exposure in prescription-claims databases

General limitations associated with the use of prescription-claims databases are that they lack information on the dosage and indication for treatment, only contain data on reimbursable medications (i.e. certain prescription and nonprescription medications are not included), and do not contain data on medications dispensed during hospital admissions, or obtained directly from the doctor (i.e. medication samples) [Crystal *et al.* 2007; Ehrenstein *et al.* 2010; Neubert *et al.* 2008].

These limitations can affect the analysis of long-term outcomes in a number of ways, many of which have been previously described. In addition, these limitations may bias the outcome under investigation. For example, Nijenhuis and colleagues investigated the long-term effects of prenatal antidepressant exposure on the development of the enteric nervous system, using dispensings for laxatives and antidiarrheal medications as a proxy for constipation and diarrhea in childhood [Nijenhuis *et al.* 2012]. The limitation here is that many laxatives and antidiarrheal medications are available without a prescription, meaning they may not be recorded in the prescription-claims database for many children. If the likelihood of women taking their child to the doctor to get a prescription for a laxative or antidiarrheal differs according to prenatal exposure status (or an additional factor such as socio-economic status), this bias could generate spurious associations between exposure and outcome.

Absence of data on nonprescription medications could introduce confound outcomes under investigation. For example, childhood asthma or wheezing has been associated with prenatal exposure to gastric acid-suppression therapy [Dehlink *et al.* 2009], and paracetamol [Rebordosa *et al.* 2008], respectively. As both of these types of medications can be obtained without a prescription, they often cannot be adjusted for when using prescription-claims databases. Notably, the potential for confounding would depend on the underlying association between these non-prescription medications and the exposure and outcome of interest.

It is also important to consider the potential for selection bias according to the way in which data on medication dispensing are recorded. An example for this can be taken from the Australian

setting. In Australia, Pharmaceutical Benefits Scheme (PBS)-claims data have been successfully linked with other population-based health datasets [Colvin *et al.* 2009]. An associated limitation with this database is the potential for underascertainment of medication exposures. Currently, medications dispensed through the PBS, where the Australian government subsidizes prescription medications for residents to ensure affordable and reliable access, are only electronically recorded if the government is required to pay a subsidy. That means data are only recorded if the cost of the dispensed medication is greater than the maximum specified cost to the consumer, known as the copayment level. Based on the PBS copayment levels for 2012, individuals with a concession card are required to pay up to AU\$5.80, while those without a concession card are required to pay up to AU\$35.40. This results in underascertainment of medication exposures in situations where medications are dispensed below the copayment level [Colvin *et al.* 2009]. Notably, this limitation can result in selection bias, where medication exposures may be differentially classified according to concession card status. In recognition of this long-standing issue, the Australian government has recently introduced legislative changes which, from 1 April 2012, require electronically recorded data to be collected on all PBS medications dispensed, including those priced below the general copayment level. In the meantime, in the situation where this is a concern, it may be necessary only to investigate outcomes amongst those with a concession card status, which reduces the potential for selection bias, but also limits generalizability of study findings.

Limitations on the generalizability of study findings also relate to other prescription-claims databases. For example, in Quebec, Canada, researchers have utilized data from the La Régie de l'Assurance Maladie du Québec drug plan [Berard *et al.* 2007]. Those eligible for, and therefore included in the drug-plan database, are recipients of social assistance (welfare beneficiaries), and workers and their families (adherents) who do not have access to a private insurance program. This is estimated to include data on 30% of women aged 15–45 years [Berard *et al.* 2007]. Therefore, studies utilizing these data are likely to underrepresent women with higher socio-economic status. While this does not threaten internal validity, it may threaten external validity if socio-economic status (or

factors associated with socio-economic status) is an effect modifier for the associations under study. As such, outcomes observed in this study cohort may not be generalizable for all women.

Medication exposure in birth-registry data

In some countries, such as Sweden, routinely collected data on medication use during pregnancy are available through birth registries. In the Swedish Medical Birth Register (SMBR), data on medication use during pregnancy have been collected since July 1994 [Kallen *et al.* 2011]. Information on first-trimester use of medications are recorded by midwives at the first antenatal visit based on maternal self-report, while information on later use during pregnancy are obtained from antenatal care records. Comparisons of data from the SMBR with data from the Prescribed Drugs Register (PDR) have demonstrated variable agreement, which is dependent on the type of medication studied [Stephansson *et al.* 2011]. Agreement ranged from 0.8% to 85.3%, and sensitivity depended on whether the medication was used for a chronic medical condition (i.e. diabetes) compared with occasional use (i.e. hayfever). For example, for antidepressants, the reported sensitivity was 59.8% (95% CI 57.7–61.9). Importantly, disagreement between the databases could be due to nonreporting or nonadherence, and for first-trimester exposure, relying on maternal self-report may be more accurate to minimize potential misclassification bias. Furthermore, the PDR has been demonstrated to provide more complete ascertainment of medication use after the first trimester than data from antenatal care [Kallen *et al.* 2011]. In this study, 38% of antidepressant use was calculated to have occurred without documentation in the antenatal care records.

Other sources of data on medication exposure

Other sources of routinely collected health data can also be used to obtain data on medication exposure during pregnancy, but these require careful validation. For example, we previously validated hospital pharmacy dispensing records as an efficient alternative to paper-based medical records in determining late-gestation exposure to antidepressants, but found that they underestimated exposure by approximately 25% [Grzeskowiak *et al.* 2010]. As a result of this underestimation it is possible that some women who were classified as unexposed and therefore allocated to the control group were actually

exposed to an antidepressant during pregnancy. This misclassification of exposure is likely to bias results towards the null. It is important to understand the limits of the data being utilized, for example, the hospital pharmacy dispensing records were only useful for identifying late-gestation exposures to antidepressants, and do not adequately identify women who were exposed to a medication during the first trimester and subsequently stopped taking it during pregnancy. As such, these records are only suitable for generating a cohort of women where exposure has occurred during late gestation [Grzeskowiak *et al.* 2010]. This, however, can still serve a useful research purpose in investigating outcomes associated with medication use during pregnancy. For example, in the case of antidepressants, fewer than 15% of women commence an antidepressant during late gestation [Colvin *et al.* 2011], women identified as exposed during late gestation are likely to have been exposed for a significant portion of the pregnancy. Therefore, late-gestation exposure is likely to reflect chronic medication use in pregnancy [Chambers *et al.* 1996; Maschi *et al.* 2008].

Losses to follow-up

One significant issue affecting traditional prospective cohort studies is the difficulty in following individuals over long periods of time. High attrition rates during a study can lead to bias if there are systematic differences between those lost and not lost to follow-up. Studies undertaken using linked data have the potential to minimize the potential for follow-up bias by reducing the need to actively follow-up individuals over time, which is costly and time consuming. In true population-based studies there are in essence no losses to follow-up instead only failures in linking records. For example, consider the use of population-based registers, such as the Danish Cancer Registry, which has had compulsory reporting of all cancers since 1987. As a result one would be confident that this dataset would contain relatively complete follow-up data for all Danish citizens. Therefore, due to the accuracy of linking records using a unique identifier, the underlying assumption made in this situation is that if an individual does not appear in the registry, they did not have a cancer-related event (i.e. relies on completeness of that data source).

The challenge presents when utilizing datasets that do not contain data for the entire population.

For example, in a previous study we utilized electronic data from child health records collected during preschool health checks [Grzeskowiak *et al.* 2012c]. From an initial cohort of 20,200 children, born between September 2000 and December 2005, follow-up data were obtained for only 7878 (39%) children. From 1997 to 2007, the period of child follow-up in this study, the average participation rate was approximately 65% of South Australian children [Gillman *et al.* 2010]. In this situation it is important to consider and explore potential explanations for loss to follow-up, and whether it is possibly due to difficulties in linking records, or is subject to bias as a result of factors associated with individuals receiving or not receiving a preschool health check.

Similarly, as previously described, claims-based databases from private health insurers (e.g. health maintenance organizations in the USA) suffer from the limitation that analyses are restricted to those who maintain coverage with the health insurer during the defined study period. The extent of this loss to follow-up is demonstrated in a study by Davis and colleagues. While investigating perinatal outcomes and congenital malformations following prenatal exposure to antidepressants, they were able to match 88% of the original maternal records to children with at least 30 days of follow-up, but just 57% of maternal records to children with at least 365 days of follow-up [Davis *et al.* 2007]. Notably, if these databases were used in an attempt to investigate child outcomes beyond 1 year, it could be expected that loss to follow-up could be even greater, potentially introducing selection bias if there are systematic differences between those who maintain their health coverage and those that change it.

Outcome ascertainment and classification

The major limitation associated with routinely collected health data is that they were never collected with the intention of being used for research purposes. Therefore, one must be clear that the data available, and the way in which outcome data are classified, are relevant in answering the research question.

In many cases it may be possible to identify outcomes of interest in inpatient or outpatient medical records by using the World Health Organization's International Classification of Diseases (ICD) codes, which is the international

standard diagnostic classification system for all diseases and other health problems. The potential exists for these codes to be used to identify a range of long-term outcomes associated with medication use during pregnancy, including the future development of diabetes, asthma, cardiovascular disease, and psychiatric diseases. However, using routinely collected health data to classify outcomes is challenging, with opportunities to both underestimate and overestimate the prevalence of the outcome under investigation. The validity of ICD codes is likely to vary across diseases, healthcare settings, and countries. As such, if these codes are used to identify outcomes, it is important to assess the correlation between coded outcomes and actual medical outcomes. For example, studies have previously evaluated the validity of using ICD codes to identify children with asthma, demonstrating that use of ICD codes alone may substantially underestimate the true prevalence of asthma. Sensitivity has been demonstrated to be as low as 24% in one study [Juhn *et al.* 2011], and 61% in another [Wakefield *et al.* 2006], depending on the algorithm used to classify outcomes. Notably, sensitivity is reported to differ according to the severity of asthma, ranging from 50% for mild asthma to 95% for severe asthma [Wakefield *et al.* 2006].

Issues can arise where classification of outcomes may differ according to additional factors associated with the exposure under investigation. For example, when relying on diagnoses made at hospital, if women taking antidepressants during pregnancy are more likely to take their children to the hospital, they may be more likely to be identified as having the outcome of interest. This type of differential misclassification bias can bias the results towards, or away from, the null. An approach to overcoming limitations of relying on ICD codes alone is to supplement this with data from prescription records. The use of an algorithm that utilizes ICD codes and prescription data has been shown to improve the identification of children with asthma, improving sensitivity from 61% to 90% [Wakefield *et al.* 2006]. Accordingly, prescription data could be used as a proxy for a range of childhood outcomes (i.e. asthma, diabetes, cardiovascular disease), but wherever possible, it is important to validate data being utilized because use of prescription records have their own limitations as previously described.

It is important to consider the difference between sensitivity and specificity of identifying outcomes using routinely collected health records. For example, Davis and colleagues investigated the risk of congenital malformations after prenatal exposure to antidepressants [Davis *et al.* 2007]. They utilized standard ICD codes to identify infants with a reported congenital malformation. Of seven infants identified as having limb malformations, case notes were obtained and reviewed for four. In one case the infant had no reported limb malformation, highlighting the potential limitations associated with relying on coded-outcome data. In contrast, Croen and colleagues investigated the association between antidepressant use during pregnancy and childhood ASD using routinely collected health data [Croen *et al.* 2011]. Prior to this study, they undertook clinical evaluations in a sample of 50 children who had an ICD code indicative of an ASD. Using the autism diagnostic interview-revised and the autism diagnostic observation schedule-generic, 94% of children met criteria for an ASD on both instruments, and 100% met criteria on at least one. While specificity is high, sensitivity could still be low. If this misclassification of outcomes is nondifferential, then it would bias results towards the null. Importantly, when relying on previously coded data to ascertain outcomes it is important to be aware of how the coding of data may have changed over time. For example, the ICD codes are approaching their 11th revision, so it is important to identify any changes that may have occurred over time, and whether these could influence outcome ascertainment, especially for studies that are comparing or including subjects over long periods of time.

Interpretation of results

A major limitation of working with routinely collected health data are that data on important confounders and/or risk factors are often missing and study findings must frequently be interpreted in their absence. Even if important pieces of data have been collected, the quality and reliability of such could be questionable unless they have been previously validated.

Precision and validity

The larger the sample size, the higher the statistical power, and the narrower the associated CI (i.e. precision). This raises a number of issues.

The first is the importance of not confusing precision with validity, as an estimate that has high precision is not necessarily one that is valid. The second is the need to be clear about separating statistically significant differences from clinically significant differences. High statistical power associated with large sample sizes makes it easier to establish even the smallest of statistically significant differences, but these differences may not be of any clinical relevance. There is also the need to be aware of the potential for finding unexpected or spurious associations, especially when working with very large datasets and looking at multiple outcomes [Sorensen *et al.* 2001].

In contrast to precision, validity is related to the confidence that can be placed on the inference drawn from a study after taking into account study methods, representativeness of the sample studied, and the nature of the population from which the sample was drawn. Notably, in some cases the use of routinely collected health data can help significantly improve external validity. Through the presence of data across a population, it is often possible to investigate outcomes amongst particularly vulnerable patient groups that would otherwise not be included in traditional prospective studies. For example, there may be many groups of women who choose not to participate in prospective studies (i.e. low socioeconomic status, mental health illness), or are commonly excluded, but these vulnerable groups could be included and studied in population-based datasets.

Confounding

As available data are limited to that which is routinely collected, studies using routinely collected health data are likely to face limitations associated with the absence of data on important confounders and/or risk factors of interest. For example, missing data on nonprescription-medication use (i.e. complementary medicines and other medicines purchased without a prescription from the pharmacy or supermarket), maternal factors (i.e. alcohol use, substance use, diet during pregnancy), child factors (i.e. diet, exercise), and other confounders and/or covariates makes it difficult to assess clearly potential associations between prenatal medication use and long-term outcomes. In particular, the inability to control for the contribution of underlying maternal disease when investigating outcomes associated with medication use during pregnancy is a constant challenge

[Grzeskowiak *et al.* 2012a]. For many outcomes, there may be the potential for significant confounding due to maternal illness (e.g. the effect of maternal depression on pregnancy outcomes when investigating the effects of antidepressant use during pregnancy) [Grzeskowiak *et al.* 2011].

Conclusion

Significant opportunities exist to utilize routinely collected health data to investigate long-term effects of medication use during pregnancy on the health of the offspring in later life. Compared with other research methodologies, the timely and efficient manner in which routinely collected health data can be used to generate large cohort studies to examine long-term outcomes is a significant advantage. With any opportunity, however, come challenges. These methodological challenges relate to aspects of study design and analysis, and include the ability to match records across datasets and over long periods of time, how medication exposures and outcomes are ascertained and classified, issues around loss to follow-up and the careful interpretation of results in light of data and study limitations. Knowledge of these challenges and recommended strategies in using routinely collected health data is essential in supporting the appropriate use and interpretation of routinely collected health data, which will contribute to improving the quality of research undertaken and ensure reliability of results obtained. Furthermore, the improved collection and evaluation of data on medication use and long-term pregnancy outcomes using routinely collected health data could enable us to strengthen safety data, encourage medication use during pregnancy when necessary, and support informed decision making based on identified benefits and risks associated with treatment.

Funding

JLM was supported by a Heart Foundation South Australian Cardiovascular Research Network Fellowship (CR10A4988).

Conflict of interest statement

The authors declare that there is no conflict of interest.

References

- Bakker, M., Kolling, P., van den Berg, P., de Walle, H. and de Jong van den Berg, L. (2008) Increase

- in use of selective serotonin reuptake inhibitors in pregnancy during the last decade, a population-based cohort study from the Netherlands. *Br J Clin Pharmacol* 65: 600–606.
- Berard, A., Ramos, E., Rey, E., Blais, L., St-Andre, M. and Oraichi, D. (2007) First trimester exposure to paroxetine and risk of cardiac malformations in infants: the importance of dosage. *Birth Defects Res B Dev Reprod Toxicol* 80: 18–27.
- Bohensky, M., Jolley, D., Sundararajan, V., Evans, S., Ibrahim, J. and Brand, C. (2011) Development and validation of reporting guidelines for studies involving data linkage. *Aust N Z J Public Health* 35: 486–489.
- Bohensky, M., Jolley, D., Sundararajan, V., Evans, S., Pilcher, D. and Scott, I. (2010) Data linkage: a powerful research tool with potential problems. *BMC Health Serv Res* 10: 346.
- Bronstein, J., Lomatsch, C., Fletcher, D., Wooten, T., Mei Lin, T., Nugent, R. *et al.* (2009) Issues and biases in matching Medicaid pregnancy episodes to vital records data: the Arkansas experience. *Matern Child Health J* 13: 250–259.
- Brook, E., Rosman, D. and Holman, C. (2008) Public good through data linkage: measuring research outputs from the Western Australian Data Linkage System. *Aust N Z J Public Health* 32: 19–23.
- Chamberlayne, R., Green, B., Barer, M., Hertzman, C., Lawrence, W. and Sheps, S. (1998) Creating a population-based linked health database: a new resource for health services research. *Can J Public Health* 89: 270–273.
- Chambers, C., Johnson, K., Dick, L., Felix, R. and Jones, K. (1996) Birth outcomes in pregnant women taking fluoxetine. *N Engl J Med* 335: 1010–1015.
- Cole, J., Ephross, S., Cosmatos, I. and Walker, A. (2007a) Paroxetine in the first trimester and the prevalence of congenital malformations. *Pharmacoepidemiol Drug Saf* 16: 1075–1085.
- Cole, J., Modell, J., Haight, B., Cosmatos, I., Stoler, J. and Walker, A. (2007b) Bupropion in pregnancy and the prevalence of congenital malformations. *Pharmacoepidemiol Drug Saf* 16: 474–484.
- Colvin, L., Slack-Smith, L., Stanley, F. and Bower, C. (2009) Pharmacovigilance in pregnancy using population-based linked datasets. *Pharmacoepidemiol Drug Saf* 18: 211–225.
- Colvin, L., Slack Smith, L., Stanley, F. and Bower, C. (2011) Dispensing patterns and pregnancy outcomes for women dispensed selective serotonin reuptake inhibitors in pregnancy. *Birth Defects Res A Clin Mol Teratol* 91: 142–152.
- Croen, L., Grether, J., Yoshida, C., Odouli, R. and Hendrick, V. (2011) Antidepressant use during pregnancy and childhood autism spectrum disorders. *Arch Gen Psychiatry* 68: 1101–1112.
- Crystal, S., Akinciqil, A., Bilder, S. and Walkup, J. (2007) Studying prescription drug use and outcomes with Medicaid claims data strengths, limitations, and strategies. *Med Care* 45: S58–S65.
- Davis, R., Rubanowice, D., McPhillips, H., Raebel, M., Andrade, S., Smith, D. *et al.* (2007) Risks of congenital malformations and perinatal events among infants exposed to antidepressant medications during pregnancy. *Pharmacoepidemiol Drug Saf* 16: 1086–1094.
- Dehlink, E., Yen, E., Leichtner, A., Hait, E. and Fiebiger, E. (2009) First evidence of a possible association between gastric acid suppression during pregnancy and childhood asthma: a population-based register study. *Clin Exp Allergy* 39: 246–253.
- Ehrenstein, V., Sorensen, H., Bakketeig, L. and Pedersen, L. (2010) Medical databases in studies of drug teratogenicity: methodological issues. *Clinical Epidemiol* 2: 37–43.
- Garbe, E., Suling, M., Kloss, S., Lindermann, C. and Schmid, U. (2011) Linkage of mother-baby pairs in the German pharmacoepidemiological research database. *Pharmacoepidemiol Drug Saf* 20: 258–264.
- Gillman, M., Oakey, H., Baghurst, P., Volkmer, R., Robinson, J. and Crowther, C. (2010) Effect of treatment of gestational diabetes mellitus on obesity in the next generation. *Diabetes Care* 33: 964–968.
- Grzeskowiak, L., Gilbert, A. and Morrison, J. (2010) Hospital pharmacy dispensing records for pharmacoepidemiology research into late gestation exposure to antidepressants. *J Pharm Pract Res* 40: 265–268.
- Grzeskowiak, L., Gilbert, A. and Morrison, J. (2011) Investigating outcomes following the use of SSRIs for treating depression in pregnancy: a focus on methodological issues. *Drug Saf* 34: 1027–1048.
- Grzeskowiak, L., Gilbert, A. and Morrison, J. (2012a). Investigating outcomes associated with medication use during pregnancy: a review of methodological challenges and observational study designs. *Reprod Toxicol* 33: 280–289.
- Grzeskowiak, L., Gilbert, A. and Morrison, J. (2012b) Exposed or not exposed? Exploring exposure classification in studies using administrative data to investigate outcomes following medication use during pregnancy. *Eur J Clin Pharmacol* 68: 459–467.
- Grzeskowiak, L., Gilbert, A. and Morrison, J. (2012c) Prenatal exposure to selective serotonin reuptake inhibitors and risk of childhood overweight. *J Dev Orig Health Dis* 3: 253–261.

- Hardy, J., Leaderer, B., Holford, T., Hall, G. and Bracken, M. (2006) Safety of medications prescribed before and during early pregnancy in a cohort of 81,975 mothers from the UK General Practice Research Database. *Pharmacoepidemiol Drug Saf* 15: 555–564.
- Herbst, A., Ulfelder, H. and Poskanzer, D. (1971) Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med* 284: 878–881.
- Holman, C., Bass, J., Rouse, I. and Hobbs, M. (1999) Population-based linkage of health records in Western Australia: development of a health services research linked database. *Aust N Z J Public Health* 23: 453–459.
- Juhn, Y., Kung, A., Voigt, R. and Johnson, S. (2011) Characterisation of children's asthma status by ICD-9 code and criteria-based medical record review. *Prim Care Respir J* 20: 79–83.
- Kallen, B. (2005) Methodological issues in the epidemiological study of the teratogenicity of drugs. *Congenit Anom (Kyoto)* 45: 44–51.
- Kallen, B., Nilsson, E. and Olausson, P. (2011) Antidepressant use during pregnancy: comparison of data obtained from a prescription register and from antenatal care records. *Eur J Clin Pharmacol* 67: 839–845.
- Kelman, C., Bass, A. and Holman, C. (2002) Research use of linked health data – a best practice protocol. *Aust N Z J Public Health* 26: 251–255.
- Kieler, H. (2010) The Nordic Health Registers – an important source when evaluating the safety of antidepressants during pregnancy. *Clin Epidemiol* 2: 205–207.
- Manning, S., Davin, C., Barfield, W., Kotelchuck, M., Clements, K., Diop, H. *et al.* (2011) Early diagnoses of autism spectrum disorders in Massachusetts birth cohorts, 2001–2005. *Pediatrics* 127: 1043–1051.
- Maschi, S., Clavenna, A., Campi, R., Schiavetti, B., Bernat, M. and Bonati, M. (2008) Neonatal outcome following pregnancy exposure to antidepressants: a prospective controlled cohort study. *BJOG* 115: 283–289.
- McMillen, I. and Robinson, J. (2005) Developmental origins of the metabolic syndrome: prediction, plasticity, and programming. *Physiol Rev* 85: 571–633.
- Meador, K., Baker, G., Browning, N., Clayton-Smith, J., Combs-Cantrell, D., Cohen, M. *et al.* (2009) Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs. *N Engl J Med* 360: 1597–1605.
- Mitchell, A. (2003) Systematic identification of drugs that cause birth defects – a new opportunity [editorial]. *N Engl J Med* 349: 2556–2559.
- Neubert, A., Sturkenboom, M., Murray, M., Verhamme, K., Nicolosi, A., Giaquinto, C. *et al.* (2008) Databases for pediatric medicine research in Europe – assessment and critical appraisal. *Pharmacoepidemiol Drug Saf* 17: 1155–1167.
- Nijenhuis, C., ter Horst, P., van Rein, N., Wilffert, B. and de Jong-van den Berg, L. (2012) Disturbed development of the enteric nervous system after *in utero* exposure of selective serotonin re-uptake inhibitors and tricyclic antidepressants. Part 2: testing the hypotheses. *Br J Clin Pharmacol* 73: 126–134.
- Olesen, C., Sondergaard, C., Thrane, N., Lauge Nielsen, G., de Jong-van den Berg, L., Olsen, J. *et al.* (2001) Do pregnant women report use of dispensed medications? *Epidemiology* 12: 497–501.
- Rebordosa, C., Kogevinas, M., Sorensen, H. and Olsen, J. (2008) Pre-natal exposure to paracetamol and risk of wheezing and asthma in children: a birth cohort study. *Int J Epidemiol* 37: 583–590.
- Sørensen, H., Johnsen, S. and Nørgård, B. (2001) Methodological issues in using prescription and other databases in pharmacoepidemiology. *Nor Epidemiol* 11: 13–18.
- Stephansson, O., Granath, F. and Svensson, T. (2011) Drug use during pregnancy in Sweden – assessed by the prescribed drug register and the medical birth register. *Clin Epidemiol* 3: 43–50.
- Wakefield, D. and Cloutier, M. (2006) Modifications to HEDIS and CSTE algorithms improve case recognition of pediatric asthma. *Pediatr Pulmonol* 41: 962–971.

Visit SAGE journals online
<http://taw.sagepub.com>

SAGE journals