Bengt Källén

Drugs During Pregnancy

Methodological Aspects

Springer

Drugs During Pregnancy

Bengt Källén

Drugs During Pregnancy

Methodological Aspects

Springer

Bengt Källén
Tornblad Institute
Lund University
Lund
Sweden

ISBN 978-3-319-40696-1 ISBN 978-3-319-40697-8 (eBook)
DOI 10.1007/978-3-319-40697-8

Library of Congress Control Number: 2016952092

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Since the middle of the 1960s, I have been working on problems associated with maternal drug use and the risk for the pregnancy and the offspring. During this time, the methodology has developed and today various strategies are used to study these complex problems. There is a very large literature in the field which is sometimes difficult to evaluate. The practical question, raised by the patient or her doctor, is if the use of a drug during pregnancy increases the risk for a negative outcome of the pregnancy and especially the risk of a congenital malformation. This question is often difficult to answer both for the patient and for her doctor, partly due to the complexity of the problem and therefore the possibility to draw wrong conclusions from published studies. This text tries to summarize my experiences and my views on these problems. It may be of some use notably for researchers who intend to enter the field.

In the text, a number of examples are given taken from the literature or based on unpublished analyses made during the more than 50 years I have been working for the National Board of Health and Welfare in Stockholm.

A limited number of references are given in each chapter. They are obviously subjectively selected from the vast literature present.

Lund, Sweden
February 2016

Bengt Källén

Contents

Abbreviations

ACE	Angiotensin converting enzyme
AGA	Birth weight adequate for gestational age
ADHD	Attention deficit hyperactivity diagnosis
ATC	Anatomical, therapeutic, chemical codes for drugs
BMI	Body mass index (weight in kg divided by squared height in meter)
d.f.	Degree of freedom
FAS	Fetal alcohol syndrome
FDA	Food and Drug Administration, USA.
HELLP	Hemolysis, elevated liver enzymes, low platelet count
ICD	International classification of diseases and related health problems
IVF	In vitro fertilization
LGA	Birth weight large for gestational age
LMP	Last menstrual period
NVP	Nausea and vomiting in pregnancy
OR	Odds ratio
PAS	Pulmonary adaption disturbances
PPHN	Persistent pulmonary hypertension of the newborn
p-value	Probability that the observed effect is random
SGA	Birth weight small for gestational age
SNRI	Serotonin and noradrenaline reuptake inhibitor antidepressant
SSRI	Selective serotonin reuptake inhibitor antidepressant
TCA	Tricyclic antidepressant
TGV	Transposition of the great vessels
TIS	Teratology information system
TTP	Time to pregnancy
95 % CI	95 % confidence interval

Introduction

1

The use of drugs during pregnancy may affect embryonic and/or fetal development. Most feared is perhaps an increase of the malformation risk. It should be realized that maternal drug use is a relatively rare cause of congenital malformations. In perhaps 25 % of such instances, genetic conditions explain the malformation, in a few percent nongenetic factors are identified, but in the remaining cases no direct explanation to the event exists. It is then easy to look for an explanation, e.g., by postulating effects of maternal drug use. As will be explained in the text, the most crucial evidence comes from epidemiological investigations. These, however, are often burdened by uncertainty, and it is easy to jump to conclusions. For the individual, risks from drug exposure are often so low that they hardly matter: if a woman has a 3 or a 4 % probability to have a malformed infant is of little significance for her, but if the drug is commonly used, even a weak effect may play a role as it can cause many malformed infants.

In the following text, examples are given from the Swedish Medical Birth Register. This register started in 1973 and was then based on documents summarizing the pregnancies. These were prepared after delivery by secretaries at the obstetric clinics (practically all births in Sweden take place in hospitals). Since 1982, data have instead been taken from copies of the original medical records which have the same format in all delivery units. With the introduction of computer-based medical records, transfer of information can be made electronically. The register contains much medical information and is supplemented with some data from Statistics Sweden. Information on factors of interest in early pregnancy (e.g., smoking, maternal weight, height, drug use) is based on interviews made by the midwives at the woman's first visit to the prenatal care (usually in week 10–12) and is thus prospective related to possible complications during pregnancy and in the neonate. Very few women do not attend prenatal care which is free of charge. Beginning on July 1, 1994, the information on drug use obtained from midwife interviews and from medical records during prenatal care was included in the register. Outcome data were obtained from the delivery records and from the pediatric examination of the newborn – all newborn infants are examined by a qualified pediatrician. Further data on

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_1

Table 1.1 A 2×2 table on exposure and outcome

	Outcome	No outcome	Total
Exposed	N_1	N_2	N_1+N_2
Unexposed	N_3	N_4	N_3+N_4
Total	N_1+N_3	N_2+N_4	$N_1+N_2+N_3+N_4$

N represents numbers

the infants born were obtained from the Hospital Discharge Register (part of the Patient Register) and also the Birth Defect Register (previously called the Register of Congenital Malformations).

A description with details of the content of the Medical Birth Register is available in http://www.socialstyrelsen.se/Publikationer2003/2003-112-3. Most presented data refer to the period 2005–2013, some to the period 1995–2013. For the period 1995–2013, there are about 1.9 million deliveries, for the period 2005–2013 about 962,000 deliveries.

Already relatively early in the text, the terms odds ratio (OR) and 95 % confidence interval (95 % CI) are used. Odds ratio is a measurement of association between exposure and outcomes. This can be illustrated with a 2×2 table (Table 1.1):

The odds for the exposed group is N_1/N_2 and the odds for the unexposed group N_3/N_4, and the odds ratio will be the quotient between the two: $(N_1/N_2)/(N_3/N_4)$. This will be an estimate of the risk of the exposure. An alternative way is to calculate a risk ratio, where the risk for the outcome in the exposed group is $N_1/(N_1+N_2)$ and the risk in the unexposed group is $N_3/(N_3+N_4)$. The odds ratio will always differ more than the risk ratio from "no effect," 1.0.

Odds ratios can be adjusted (adjusted odds ratios) for the possible influence of other factors in ways which will be discussed later in the book. The 95 % confidence interval (95 % CI) shows the likely interval within which the true OR or RR lies.

The "Alert Clinician"

2

Since the thalidomide tragedy around 1960, when this new sleeping drug showed to be a strong human teratogen, the question about the safety of drug use during pregnancy has been of great interest, not only for researchers but also for administrators, doctors who prescribe drugs, and their patients. A very large number of scientific papers are published each year on this subject, sometimes finding no evidence of harm, sometimes pointing out possible risks for the unborn child. Continuously, efforts are made to summarize and evaluate such data and to formulate recommendations regarding drug use during pregnancy.

Before thalidomide, rather little interest was paid to the question of drug teratogenesis. This can be illustrated by the first international conference on congenital malformations, held in London in 1960, 1 year before the detection of the strong teratogenic properties of thalidomide. One of the grand old men in teratology, Joseph Warkany, gave a lecture on environmental teratogenic factors. Among man-made such factors, Warkany mentioned antimetabolites and synthetic progestins. The former drugs had been used in experimental teratology in order to cause malformations in laboratory animals, and the use of aminopterin as an abortificant had shown that such drugs could cause serious malformations also in the human embryo. This was to be expected from the mode of action of these drugs – they were used to kill rapidly growing cancer cells and could therefore be expected to damage also the rapidly dividing embryonic cells. Synthetic progestins could sometimes disturb the development of female fetal genital organs which could be masculinized – this again was easy to understand as the development of the male genital organs is normally stimulated by androgen, and synthetic progestins sometimes have androgenic effects.

Also other drug categories had been used in experimental teratology to cause malformations in laboratory animals, but usually high doses were needed, and the significance of these for the human situation was usually regarded as dubious.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_2

2.1 Thalidomide

Then thalidomide arrived. This tragedy has been described repeatedly in the literature. Briefly, in 1956 the German drug industry Grünenthal marketed a new drug containing thalidomide and with a wide range of indications, notably influenza – the drug was first sold under the name Grippex. In 1957 the same substance was marketed under the name of Contergan, and the indication for use was mainly as a sedative but it could also be used at nausea and vomiting in pregnancy (NVP). The first known case of a thalidomide malformation in man was born in December 1956, the second in March 1957, and then there was an increasing number during 1958, culminating in 1961 (Lenz 1988). The epidemic of infants with severe limb malformations in Germany had been noted and was published in September 1961 and was called the Wiedemann dysmelia syndrome (1961), but no cause was initially identified.

A link with thalidomide was not published until December 16, 1961, by McBride (1961) in Australia and was suggested at the same time by Lenz (1961) in Germany. Already on December 2, 1961, the British firm Distillers Company had reported on a suspicion that thalidomide could cause malformations and their drugs containing this substance were withdrawn from the market in England. The McBride report was a short letter stating that 20 % of infants born of women who had used thalidomide during pregnancy were malformed, but no actual numbers were given. According to Fraser (1988), the observation was based on six malformed cases, the first one born in May 1961.

The finding that a sedative with few side effects in the adult could cause severe malformations made the scientific world to reconsider the importance of maternal drug use for the origin of malformations, and the medical journals received many reports on maternal drug use followed by the birth of a malformed infant. Some of these resulted in the identification of teratogenicity. The history of some of them will be summarized.

2.2 Anticonvulsants

An early concern referred to the use of anticonvulsants during pregnancy. The first study on this subject was probably that by Janz and Fuchs (1964). These authors studied the outcome of 262 deliveries when the mothers had used anticonvulsants and found that only five infants were malformed (2.2 %). Their conclusion was that anticonvulsant therapy had no teratogenic effect, but they did notice that three of the five malformed infants had cleft lip.

After having observed six infants with cleft lip/palate born by women with epilepsy, Meadow in 1968 asked in a letter to Lancet about further such cases born in England and collected in this way a total of 32 cases (Meadow 1970). Many of them had other defects, notably cardiovascular malformations. This number seemed to be higher than what could be expected to occur by chance. The author was careful to stress that this observation was no proof of causality, but he invited further studies.

Numerous such studies have been published since then, and the suggestion by Meadow has been amply verified. Most anticonvulsants have been shown to have teratogenic properties, and among the typical malformations seen after the use of the "old" drugs, for instance, phenytoin, were orofacial clefts. An excellent summary of the early literature was published by Bossi (1983).

A new aspect on the teratogenicity of anticonvulsants came from French investigators who published a letter in Lancet, demonstrating an association between maternal use of an anticonvulsant, valproic acid, and infant spina bifida (Robert and Guibaud 1982). These authors described that among 72 infants with lumbosacral neural tube defects, nine had an epileptic mother who had taken valproic acid during pregnancy. An earlier report, based on animal experiments, had identified valproic acid as a stronger animal teratogen than other anticonvulsants (Brown et al. 1980). An international study (Bjerkedal et al. 1982) of maternal anticonvulsant use and infant spina bifida found that the association with valproic acid seemed to be much stronger than that with other anticonvulsants. The strong teratogenic effect of valproic acid, not restricted to spina bifida, has been demonstrated in numerous later studies with various designs, and valproic acid is nowadays one of the best known relatively strong teratogens with a rather wide spectrum of malformation outcomes and also serious effects on the long-term development of the exposed child.

2.3 Warfarin

The first report which suggested a typical malformation in infants whose mothers had used warfarin came in 1966 (di Saia 1966). It was followed by further reports, e.g., that by Kerber et al. (1968). The typical abnormalities were nasal hypoplasia and a skeletal condition resembling the genetic condition chondrodystrophia calcificans. According to a multicenter prospective study, the use of coumarin drugs in the first trimester resulted in a nearly fourfold increase in the rate of major congenital malformations, but the rate of typical warfarin embryopathy was low (Schaefer et al. 2006).

2.4 Drugs for Thyrotoxicosis

The first observations on a possible teratogenicity of some drugs used at thyrotoxicosis were made in the 1970s (Milham and Elledge 1972; Mujtaba and Burrow 1975) when an association between maternal use of methimazole and scalp defects was noted. Later cases were reported with such or other defects, and a methimazole embryopathy syndrome has been delineated (Clementi et al. 1999) in which some serious malformations are part, e.g., choanal atresia and esophageal atresia. Propylthiouracil is generally regarded as non-teratogenic, while the evidence for a teratogenic effect of methimazole and related substances seems rather well established. An exception is a Danish study which found an increased malformation risk

also after propylthiouracil but with other specific malformations than those noted after methimazole, mainly rather minor and variably recorded conditions (Andersen et al. 2014)

2.5 Retinoids

A classical method to cause malformations in rats is to give high doses of vitamin A. Derivatives of vitamin A, retinoids, have been used in medical practice for various indications. In a letter to the Lancet in 1983, Rosa pointed out that two such drugs were used, oral isotretinoin for severe acne (introduced in USA in 1982) and etretinate for psoriasis. Rosa reported that five malformed infants (four of them with hydrocephaly) had been reported by the manufacturer of isotretinoin to the Food and Drug Administration. Lammer et al. (1985) collected and analyzed adverse drug reaction reports from the manufacturer, from the Food and Drug Administration and from the Center of Disease Control. They investigated 154 infants, exposed to isotretinoin in early pregnancy. A total of 21 malformed infants were described. Cranium and face were affected in 17 of them, heart in 12, central nervous system in 18, and thymus in seven.

Etretinate has been linked to a high risk for severe malformations, notably exencephaly but also skeletal malformations. An early report (Happle et al. 1984) described 19 women who had used etretinate during pregnancy. Three infants had skeletal defects, one spontaneously aborted fetus had a meningomyelocele, and two fetuses which were studied after induced abortions had cerebral abnormalities. Ten infants were normal and so were three fetuses after induced abortions. Acitretin is metabolized to etretinate and has apparently similar teratological properties (Barbero et al. 2004) but a shorter half-life.

2.6 Mycophenolate Mofetil

This drug, an immunosuppressant which is mainly used after organ transplantation, is teratogenic in animal experiments, and some publications have linked its use to an increased risk for congenital malformations. No epidemiologic investigation has explored this relationship, but a typical syndrome has appeared (Peres-Ayetes et al. 2008; Vente et al. 2008): cleft lip/palate, microtia, and external auditory canal atresia. Other malformations were also present including limb defects. The validation of this syndrome needs further data and notably an estimate of absolute risk. An existing evaluation of a 27 % risk was based on four malformed among 15 exposed infants, reported to the US transplantation register (Sifontis et al. 2006). The 95 % confidence interval of this estimate is 8–55 %.

The above summaries of the history of the identification of some teratogenic drugs illustrate the significance of what has been called "the alert clinician," the clinical observation of the association between malformations in the newborn and drugs used by the mother. Many of these observations refer to rare or complex

conditions – it seems to be more likely to suggest an association between maternal use of a specific drug and an unusual malformation than with a common malformation. The frequency of random associations between drug and malformation will be higher in the latter than in the former situation. Obviously many such observations have been made which have not resulted in the identification of teratogenicity, and a single observation has its main value for the initiation of further studies. For such studies, epidemiological methods are used as will be discussed in this book.

References

Andersen SL, Olsen J, Wu CB, Laurberg P (2014) Severity of birth defects after propylthiouracil exposure in early pregnancy. Thyroid 24:1533–1539

Barbero P, Loterszlein V, Bronberg R, Perez M, Alba I (2004) Acitretin embryopathy: a case report. Birth Defects Res Part A 70:831–833

Bjerkedal T, Czeizel A, Goujard J, Källén B, Mastroiacovo P, Nevin N et al (1982) Valproic acid and spina bifida. Lancet 2:1096

Bossi L (1983) Fetal effects of anticonvulsants. In: Morselli PL, Pippenger CE, Penry JK (eds) Antiepileptic drug therapy in pediatrics. New York, Raven Press

Brown NA, Kao J, Fabro S (1980) Teratogenic potential of valproic acid. Lancet 1:660–661

Clementi M, Di Gianantonio E, Pelo E, Mammi I, Basile RT, Tenconi R (1999) Methimazole embryopathy: delineation of the phenotype. Am J Med Genet 83:43–46

Disaia PJ (1966) Pregnancy and delivery of a patient with a Starr-Edwards mitral valve prosthesis. Obstet Gynecol 28:469–472

Fraser FC (1988) Thalidomide retrospective: what did we learn? Teratology 38:201–203

Happle R, Traube H, Bounameaux Y, Fisch T (1984) Teratogen Wirkung von Etretinat beim Menschen. Dtsch med Wochenschr 109:1476–1480

Janz D, Fuchs U (1964) Sind antiepileptische Medikamente während der Schwangerschaft schädlich? Deutsch Med Wochenschrift 89:241–244

Kerber IJ, Warr OD 3rd, Richardson C (1968) Pregnancy in a patient with a prosthetic mitral valve. Associated with a foetal anomaly attributed to warfarin sodium. JAMA 203:223–225

Lammer EJ, Chen DT, Hoar RM, Agnish ND, Benke PJ, Braun JT et al (1985) Retinoic acid embryopathy. N Engl J Med 313:837–841

Lenz W (1961) Kindlische Missbildungen nach Medikament-Einnahme während der Gravidität? Dtsch Med Wochenschrift 86:2555–2556

Lenz W (1988) A short history of thalidomide embryopathy. Teratology 38:203–215

McBride WG (1961) Thalidomide and congenital abnormalities. Lancet 2:1358

Meadow SR (1970) Congenital abnormalities and anticonvulsant drugs. Proc Roy Soc Med 63:12–13

Milham S, Elledge W (1972) Maternal methimazole and congenital defects in children. Teratology 5:125

Mujtaba Q, Burrow GN (1975) Treatment of hyperthyroidism in pregnancy with propylthiouracil and methimazole. Obstet Gynecol 46:282–286

Peres-Ayetes A, Ledo A, Bosso V, Sáenz P, Roma E, Poveda JL et al (2008) In utero exposure to mycophenolate mofetil. A characteristic phenotype. Am J Med Genet A 146A:1–7

Robert E, Guibaud P (1982) Maternal valproic acid and congenital neural tube defects. Lancet 2:937

Rosa F (1983) Teratogenicity of isotretinoin. Lancet 2:513

Schaefer C, Hanneman D, Meister R, Eléfant E, Paulus W, Vial T et al (2006) Vitamin K antagonists and pregnancy outcome. A multi-centre prospective study. Thromb Haemost 95:949–957

Sifontis NM, Coscia LA, Constantinescu S, Lavelanet AF, Moritz MJ, Armeni VT (2006) Pregnancy outcomes in solid organ transplant recipients with exposure to mycophenolate mofetil and sirolimus. Tansplantation 82:1698–1702

Vente M, Perez AA, Ledo A, Boso V, Carey JC (2008) Mycophenolate mofetil during pregnancy: some words of caution. Pediatrics 122:184–185

Warkany J. Environmental teratogenic factors. First International Conference of congenital malformations. Papers and discussions presented at the First International Conference of congenital malformations, London 18–22 July 1960. JB Lippincott Company, Philadelphia and Montreal, 1961

Wiedemann H-R (1961) Hinweise aut eine derzeitige Haufung hypo- und aplastischer Fehlbildungen der Gliedmassen. Med Welt 37:1863–1866

Animal Experiments and Adverse Drug Reaction Reports

3

3.1 Animal Experiments

The history of teratology experiments with laboratory animals goes back to the early twentieth century. They should be looked upon as tools to disturb the normal development in order to get a better understanding of the forces which control embryogenesis. Much of the early studies referred to the effect of hypovitaminosis, beginning with the studies on pigs given an A-vitamin-deficient diet which resulted in anophthalmia in the offspring (Hale 1933). In the 1940s, many studies were published on the effect of antimetabolites which could cause malformations.

Baxter and Fraser (1950) demonstrated that treatment of pregnant mice with cortisone could induce cleft palate, and similar effects have been found in some other – but not all – species studied. A strong variability in sensitivity between different inbred mouse strains was demonstrated by Kalter (1965), 12 % in the CBA and 100 % in the A/J strain. Extensive studies of this model have been made during the decades, and also other types of corticosteroids have been tested. The relatively constant findings have resulted in a belief that the same should be true also for man, and many studies have tried to investigate whether the use of corticosteroids during early pregnancy increases the risk of orofacial clefts and notably of median cleft palate in man. The results of epidemiological studies have varied, but it is possible that a slight over-risk exists but it actually seems to be valid mostly for cleft lip/palate and not for cleft palate. Anyway, there is a clear contrast between the relatively constant findings in the animal experiments and the uncertain relationship suggested in man. The animal data, however, for a long time resulted in a warning for the use of corticosteroids during pregnancy, also when they were used as an inhaled drug at asthma, a situation on which large amounts of data are now available without any certain signs of a teratogenic effect. The odds ratio for any major malformation after maternal use of inhaled glucocorticoids in early pregnancy was 1.03 (95 % CI 0.94–1.14) (Källén 2009).

Another drug which belongs to the classical animal teratogens is acetylsalicylic acid which caused cleft lip and other defects in mice (Trasler 1965) and is also

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_3

teratogenic in other species. In spite of the fact that single cases of gross malformations have been associated with maternal use of aspirin, proper epidemiological studies have in most cases found no such effect. An example is a prospective study with 8371 exposed infants where the odds ratio for a congenital malformation was 0.94 (95 % CI 0.83–1.06) (Källén 2009).

On the other hand, when thalidomide was identified as a strong teratogen, it took about half a year until anyone succeeded in producing a similar malformation with the drug in animals, white New Zealand rabbits (Somers 1962). Further studies have shown that primates could react in the same way as humans but that results of experiments with rodents and other common laboratory animals gave unclear results. Primates seem to be remarkably sensitive to the teratogenicity of this drug, and it is doubtful if the present protocols for preclinical testing of drugs for teratogenicity would have identified the risk associated with thalidomide.

After the thalidomide tragedy, various actions were taken to prevent a repetition. One was that it became compulsory to test new drugs in animal reproduction experiments, and such experiments were to a large extent standardized. Such tests try to identify reproductive toxicity effects which are judged likely to be valid also for humans, and drugs with such properties will not be brought to clinical testing and marketing. There will thus be no way to evaluate how predictive the tests are for the human situation. Very few drugs with a strong human teratogenicity have been identified after thalidomide – this may be the result of a high effectiveness of the preclinical testing or just an effect of the fact that such drugs actually are rare. Some drugs with positive animal tests are brought to the market, but their use in pregnant women is discouraged in order to be on the safe side. Some exposures will occur nevertheless, but obviously a moderately strong teratogenicity will be difficult to detect because of a low statistical power.

Relatively few drugs are demonstrated as definitely teratogenic for the human embryo. As pointed out above, it took some time before the teratogenic property of thalidomide (identified by alert clinicians) could be demonstrated, and the best animal models use primates, hardly suitable for routine screening procedures. Another example of a human teratogen is methimazole with a typical teratogenic activity (see above). No teratogenic activity was detected in mice or rats even at high dosage, and a preclinical testing had probably not identified the human teratogenic effect (Mallela et al. 2014). The quoted paper tended to draw the conclusion that methimazole lacked teratogenic properties; a more likely interpretation is perhaps that mice and rats are not suitable animals for a study of this problem. In other circumstances, animal studies identified risks which were later verified in human studies; the best example is perhaps isotretinoin. As mentioned above, high doses of vitamin A have long been used in experimental teratology and can cause CNS malformations, notably exencephaly, but also, for instance, eye and limb defects and facial clefts. Among different retinoids, all-trans retinoic acid seems to be a stronger teratogen than 13-cis-retinoic acid (isotretinoin), and it is thought that the teratogenic activity of the latter is the result of its metabolism to all-trans retinoic acid, a process which differs between species (Adams 1993). The high teratogenic activity in some species (man and other primates) may also be due to slow elimination and

metabolism and high placental transfer (Nau 2001). In order to mimic the human teratogenic process, experiments on other primates are needed. Actually, the teratogenic effect of high doses of vitamin A (for instance, in the mouse and rat) may not be related to the rather specific teratogenicity of isotretinoin in man.

Animal studies are useful in order to clarify the mechanism of a human teratogenic effect, seen with a certain drug. For this purpose, a suitable species or strain should be identified reacting in a similar way as man does. Experiments in rabbits and other animal species suggested various explanations of the teratogenic activity of thalidomide.

Another example is the very strong association between maternal use of valproic acid and hypospadias in male offspring. Male genital phenotype is due to an androgenic stimulation of the genital rudiments, and hypospadias can be regarded as the result of an incomplete masculinization of the genital organs. As expected, hypospadias can be obtained in animals with antiandrogenic drugs. Even high doses of valproic acid did not affect genital development in rats in experiments where an antiandrogenic drug (flutamide) did (Källen 2004). The effect of valproic acid most likely has another explanation than an antiandrogenic effect.

3.2 Adverse Drug Reaction Reports

A second result of the thalidomide tragedy was an intensification of the reporting of suspected adverse reactions of drugs, including observations of congenital malformations or other adverse effects in infants whose mothers had used drugs. Such information is collected by the World Health Organization, at present by the branch in Uppsala, Sweden. Obviously reproduction anomalies are only a small part of all adverse reactions reported. I have some experience of this system because reports in Sweden on suspected adverse reactions associated with congenital malformations or other reproductive abnormalities have to some extent been referred to me for evaluation.

There are different questions which can be raised by such a report. In some instances, the reported association is known from the literature; sometimes it is not. Given the large number of drugs used and the large number of different reproductive abnormalities which can occur, many random associations between drug use and outcome are to be expected. In many cases, this can be made likely, for instance, if the exposure has occurred too late to be able to have caused the reported malformation (see Chap. 12). In other cases, there is a possibility that a causal association exists. It *could* be a case of an "alert clinician" noting the first known teratogenic effect of a drug. The probability for an association increases, of course, if independent reports on a specific association are obtained. All such reports should therefore be stored for reference purposes. A rather high level of detail is needed in the report, notably with respect to the outcome, for instance, a detailed diagnosis of the malformation involved. Unfortunately, such details are often missing, and one has put the malformations into groups like "limb malformations" or "musculoskeletal malformations." If used correctly, this is an effective way to get an indication that a new

association between maternal drug use and infant abnormality has appeared which should initiate a study of the association in an independent material.

When the report refers to an association which is already known from the literature, another aspect may be relevant, the question of causality in the individual case. If the association is very strong, e.g., the association between use of valproic acid and spina bifida or that between methimazole and choanal atresia, the question of causality is relatively simple. If the risk increase is 20 times, 19 out of 20 exposed infants with that malformation will be caused by the drug and only one will be coincidental. If, on the other hand, the risk is only doubled, half of the associations between exposure and outcome will be random and half causal. If the risk increase is lower than that, it is more likely that an individual case is random than causal.

It does not matter if it is a rare malformation or a common malformation – the important point is the strength of the association. In the evaluation of an individual case, one also has to consider the exposure time. For example, if the drug exposure occurred during the organogenesis of the malformation in question, it is compatible with causality, if it occurred outside the period of organogenesis it speaks for randomness. Later in this book, we will come back to the problem of timing and risk estimates which are not always as simple as they may seem (Chap. 12).

There are problems in the individual case with statements that a specific drug has caused a specific abnormality. First, as just stated, it is always a question of probabilities. Second, if a drug is pointed out as the cause of an abnormality, other possible etiological factors may be ignored, e.g., a genetic risk. If an infant's microcephaly is blamed on maternal use of a drug and it really is an autosomal recessive condition, the couple may disregard the recurrence risk of 25 % and just avoid the pinpointed drug.

In Table 3.1, 27 adverse reports are summarized which I have evaluated the last year or so. It can be noted that ten of them refer to antidepressants. The use of such drugs during pregnancy is relatively common and associated with some problems related to the offspring. The risk for a congenital malformation is usually low, perhaps higher with a tricyclic antidepressant (TCA) than with a selective serotonin reuptake inhibitor (SSRI). When used in the later part of the pregnancy, temporary effects on the neonate are common, e.g., respiratory problems.

There are rather good evidence that at least use of SSRI increases the risk of a rare but serious complication, persistent pulmonary hypertension of the newborn, PPHN. In a study from Sweden (Reis and Källén 2010), a risk increase of 2–3 times was found in infants born after week 33. Among the cases listed in the table, four refer to PPHN, two of them after exposure to sertraline (an SSRI) and two after exposure to venlafaxine (an SNRI). The two cases with sertraline exposure and PPHN were more likely caused by the drug than they were coincidental. An association between venlafaxine use and PPHN has never been proved, but due to the low rate of this complication, a very large number of exposures are needed to detect an association and causality in the reported cases cannot be dismissed. There is a need for a large-scale study of maternal use of venlafaxine.

Long-term effects on the development of the child may exist but are difficult to demonstrate and interpret (Källén et al. 2013). The association between use of paroxetine

Table 3.1 Summary of some recently evaluated adverse drug reports concerning reproduction outcome

Number	Drug	Outcome	Known association
1	Venlafaxine	PPHN	No
2	Venlafaxine	PPHN	No
3	Sertraline	PPHN	Yes
4	Sertraline	PPHN	Yes
5	Sertraline	Infant liver affection	No
6	Sertraline	Neonatal respiratory problem	Yes
7	Citalopram	Megaloureter/hydronephrosis	No
8	Citalopram	Positional foot defect (2 cases)	No
9	Paroxetine, fluoxetine	Late language development	No
10	Duloxetine, fluconazole, oxazepam	Miscarriage	No
11	Valproic acid	Epilepsy, ADHD, developmental delay	Yes
12	Valproic acid	Late language development	Yes
13	Lamotrigine	TGV	No
14	Dixyrazine	ADHD and epilepsy	No
15	Perphenazine	Autism	No
16	Oxycodone, diclofenac, ondansetron	Cardiac defect	Yes
17	Misoprostol	Neonatal asphyxia	No
18	Misoprostol	Anencephaly	No
19	Magnesium before delivery	PAS, hypotonia	Yes
20	Mesalazine	TGV	No
21	Nitrofurantoin	Severe ear malformation	No
22	Fingolimod	Lissencephaly	No
23	Glatiramer	Severe brain malformation	No
24	Adalimumab	Severe cardiac defect	No
25	Vaccination against H1N1 influenza	Autism	No
26	Vaccination against H1N1 influenza	Autism	No
27	Mercaptopurine (paternal exposure)	Schizencephaly	No

ADHD attention deficit and hyperactivity diagnosis, *PAS* pulmonary adaption disturbances, *PPHN* persistent pulmonary hypertension of the newborn, *TGV* transposition of great vessels

and fluoxetine and delayed language development in the infant agrees with the results of a study by Skurtveit et al. (2014) which found an effect on language development from long-term use of SSRI during pregnancy. In the present case, exposure occurred only during months 2–3, and according to Skurveit et al. short-term use was not found to affect speech development, which speaks against causality in the reported case.

Three cases refer to use of citalopram and relatively minor and common malformations. No certain teratogenicity of citalopram has been demonstrated even though an association with cardiac septum defects has been suggested in some studies but not in other. It is likely that these cases represent coincidental associations.

Two cases refer to exposure to valproic acid and long-term effects on child development. Such effects are known to occur (Banach et al. 2010) and causality is likely. Two other cases (#14, 15) refer to long-term effects on child development from psychoactive drugs. Large studies of maternal use of such drugs and autism, ADHD, and other developmental deviations in the offspring are badly needed but offer large methodological difficulties.

One case represents the association of maternal use of lamotrigine and infant TGV. Many anticonvulsants increase the risk of cardiovascular defects, but data for lamotrigine indicate no such strong association but more information is needed. To detect a three times increase in the rate of TGV (supposing a base rate of 1/2000), one would need information on about 5800 exposures and compare them with a large control material. A similar situation exists for report #20. Mesalazine has been associated with a moderate teratogenic risk, notably cardiovascular defects. In the largest available study (Källén 2014), only 2050 exposures were included, less than half of what would be needed to have a reasonable chance to demonstrate a significant association with TGV.

The case reporting the use of misoprostol and infant anencephaly is interesting (#18). Misoprostol, when used to induce a miscarriage, has been linked to some congenital malformations but as far as I know not to anencephaly. In this case, exposure occurred in week 6 which is too late to cause anencephaly (under the assumption that dating was correct).

Case #16 represents a complex exposure situation. The nature of the cardiac defect was not known, but the child was operated upon for its malformation. Among the mentioned drugs, notably ondansetron has been linked to congenital malformations in the offspring and notably to cardiovascular defects (Danielsson et al. 2014), but this finding mainly refers to septal defects. The total risk increase for a cardiovascular defect in that study was 1.6 which means that it is more likely that the association in the reported case was random than causal.

Two adverse reports (#23, 24) refer to drugs used for multiple sclerosis and severe but different brain malformations. One infant, whose mother had previously used fingolimod, had lissencephaly; the other infant whose mother had used glatiramer during pregnancy had a frontal encephalocele, agenesis of corpus callosum, and a suspected optic nerve hypoplasia. Use of such drugs during pregnancy is very rare, and both drugs are contraindicated during pregnancy. The first exposure to glatiramer in Sweden was registered in 2003 – up to the end of 2013 there were only 39 exposures registered. Among these, three had a malformation diagnosis: one had an atrium septum defect, one a larynx abnormality, and one an unstable hip. No significant teratogenicity was thus seen in this small material – but it would take a long time to collect enough data to evaluate the reported association.

No woman giving birth had reported the use of fingolimod during early pregnancy up to 2013. In the reported case, the woman had stopped using the drug about

2 years before the pregnancy, and even if the drug has a long half-life, it seems improbable that it could cause a malformation such a long time after stopping the drug. The case is complicated by the fact that the woman had a previous pregnancy with fingolimod when a cardiac malformation (tetralogy of Fallot) was detected which resulted in an induced abortion. There is very little information on the use of fingolimod during pregnancy – a study of 28 live births showed two malformed ones, one with anencephaly, the second with a bowing of the tibia (Karlsson et al. 2014). The drug is contraindicated during pregnancy mainly because of animal data.

The association between maternal use of nitrofurantoin and infant ear malformation (#21) is probably random. No teratogenic effect of adalimumab (#24) has been demonstrated. The first exposure during pregnancy in Sweden occurred in 2008, and up to and including 2013, 55 exposures in early pregnancy are known. Two of the infants had an anomaly; one had a hydronephrosis and one a tongue tie. To get a better idea of possible risks, one would need many more exposed cases.

Two reports (#25 and 26) concern autism in children born by women who were vaccinated against H1N1 influenza during pregnancy. These reports cannot at present be evaluated because there are no studies on the subject available. It would be possible to follow a large number of infants after such maternal vaccinations in order to identify autism cases as information on such vaccinations during pregnancy is available (Källén and Olausson 2012).

The last case in the table (#27) refers to paternal exposure for mercaptopurine. We will discuss the risks with paternal exposures later on in the book. A possible effect could be an increased mutation rate caused by the drug. This could result in a congenital malformation caused by a dominant gene. In the literature, only small studies are available of male mercaptopurine exposure (Hoeltzenbein et al. 2012). The genetic background of the observed malformation (schizencephaly) is unclear, and a dominant mutation cannot be excluded.

These examples illustrate the complexity in the evaluation of reported suspected adverse reactions of drug use in connection with pregnancy. It is often not possible to state if the association between drug use and outcome was causal or not. The important thing is to collect the data and search for the repeated occurrence of associations between specific drugs and specific malformations and to follow-up such observations with proper epidemiological studies, when possible.

References

Adams J (1993) Structure-activity and dose–response relationships in the neural and behavioral teratogenesis of retinoids. Neurotox Teratol 15:193–202

Banach R, Boscovic R, Einarson T, Koren G (2010) Long-term development outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs a meta-analysis of cohorts studied. Drug Saf 33:73–79

Baxter H, Fraser FC (1950) Production of congenital defects in offspring of female mice treated with cortisone. McGill Med J 19:245–249

Danielsson B, Norstedt-Wikner B, Källén B (2014) Use of ondansetron during pregnancy and congenital malformations. Reprod Toxicol 50:134–137

Hale F (1933) Pigs born without eyeballs. J Hered 24:105–106
Hoeltzenbein M, Weber-Schoendorfer C, Borisch C, Allignol A, Meister R, Schaefer C (2012) Pregnancy outcome after paternal exposure to azathioprine/6-mercaptopurine. Reprod Toxicol 14:364–369
Källen B (2004) Valproic acid is known to cause hypospadias in man but does not reduce anogenital distance or causes hypospadias in rats. Basic Clin Pharmacol Toxicol 94:51–54
Källén B (2009) Drugs during pregnancy. Nova Biomedical Books, New York
Källén B (2014) Maternal use of 5-aminosalicylates in early pregnancy and congenital malformations in the offspring. Scand J Gastroenterol 49:442–448
Källén B, Olausson PO (2012) Vaccination against H1N1 influenza during pregnancy and delivery outcome: a Swedish register study. BJOG 119:1583–1590. doi:10.1111/j.1471-0528.2012.03470.x
Källén B, Borg N, Reis M (2013) The use of central nervous system active drugs during pregnancy. Pharmacoceuticals 6:1221–1286
Kalter H (1965) Experimental investigation of teratogenic action. Ann N Y Acad Sci 123:287–294
Karlsson G, Francis G, Koren G, Heining P, Zhang X, Cohen JA et al (2014) Pregnancy outcomes in the clinical development program of fingolimod in multiple sclerosis. Neurology 82:674–680
Mallela MK, Strobi M, Poulsen RR, Wendler CC, Booth CJ, Rivkees SA (2014) Evaluation of developmental toxicity of propylthiouracil and methimazole. Birth Def Res (Part B) 101:300–307
Nau H (2001) Teratogenicity of isotretinoin revisited: species variation and the role of all-trans-retinoic acid. J Am Acad Dermatol 45:183–187
Reis M, Källén B (2010) Delivery outcome after maternal use of antidepressant drugs in pregnancy: an update using Swedish data. Psychol Med 10:1723–1733
Skurtveit S, Selmer R, Roth C, Hernandez-Dias S, Handal M (2014) Prenatal exposure to antidepressants and language competence at age three: results from a large population-based pregnancy cohort in Norway. BJOG 121:1621–1
Somers GF (1962) Thalidomide and congenital abnormalities. Lancet 2:912–913
Trasler DG (1965) Aspirin-induced cleft lip and other malformations in mice. Lancet 1:606

Some Epidemiological Principles

4

In studies on the possible relationship between maternal use of a drug and an infant outcome like a congenital malformation, the first epidemiological question to answer is: do maternal use of this drug and the occurrence of the malformation in the infant occur together more often than expected?

4.1 Population Studies

A crucial point is of course to estimate how often the outcome (e.g., a malformation) occurs in the absence of maternal use of the drug. If we have information on drug use by all women and presence of malformations in all infants, this is a rather simple problem, as seen in Fig. 4.1. The population (all pregnant women) is represented by a square and is divided into two vertical areas: one which represents the women who used the drug (exposed) and the other women who did not use the drug. Another division occurs according to the presence of the adverse reproduction outcome (e.g., malformations): one horizontal area with malformations and another without malformations. The striped area represents malformed infants, exposed to the drug, under the assumption that no relationship exists, that is, the expected number of such outcomes. If the rate of malformations is p among all pregnancies and N is the number of exposed pregnancies, the expected number of malformations among the latter will of course be p*N.

Nowadays there are areas or countries where it is possible to study the whole population of pregnant or delivered women by using regional or national health registers. These possibilities are relatively recent, and before that time, other ways to estimate the expected number had to be used, mainly case-control or cohort studies. Both techniques are based on sampling techniques.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_4

Fig. 4.1 Diagram illustrating exposure and adverse outcome in the population

4.2 Case-Control Studies

A case-control or case-referent study is based on the collection of exposure data for cases (e.g., malformed infants) and controls (e.g., non-malformed infants) and then the exposure rates in the two groups are compared. This is illustrated in Fig. 4.2. Among all non-malformed infants in the population, a group is identified one way or another, and information on maternal drug use among them is obtained. This control group is thus a sample of the non-malformed individuals and, if representative, will give an estimate of exposure rate among mothers of non-malformed infants.

In most instances, information on drug use is obtained from questionnaires or interviews, rarely from medical records produced already during pregnancy. The former exposure information is thus retrospective which carries problems because of recall or interviewer bias. This phenomenon will be discussed later on (Chap. 7).

In some large-scale studies, e.g., those from the US National Birth Defects Prevention Study (Yoon et al. 2001) or the Slone Epidemiology Center Birth Defects Studies (Yao et al. 2013), exposure data were collected by telephone interviews a considerable time after birth which may increase the risk for recall bias. In this situation a further problem exists. One usually has a nonresponse rate of about 30 %. The risk for a selective nonparticipation is large which can give false results. It can be debated if retrospective case-control studies concerning maternal drug use and infant outcome should at all be performed when other and more reliable methodologies exist. The technique may have a place in studies of factors which are more difficult to identify in an objective way, like nutritional factors (e.g., Botto et al. 2015) or the effect of hot water baths – but the same skepticism should be kept in the evaluation of the results.

We can illustrate the problems with these studies with data published on the effect of maternal use of opioid analgesics and infant congenital malformations

Fig. 4.2 Diagram illustrating the principle of a case-control study. The sizes of the two hatched areas are compared

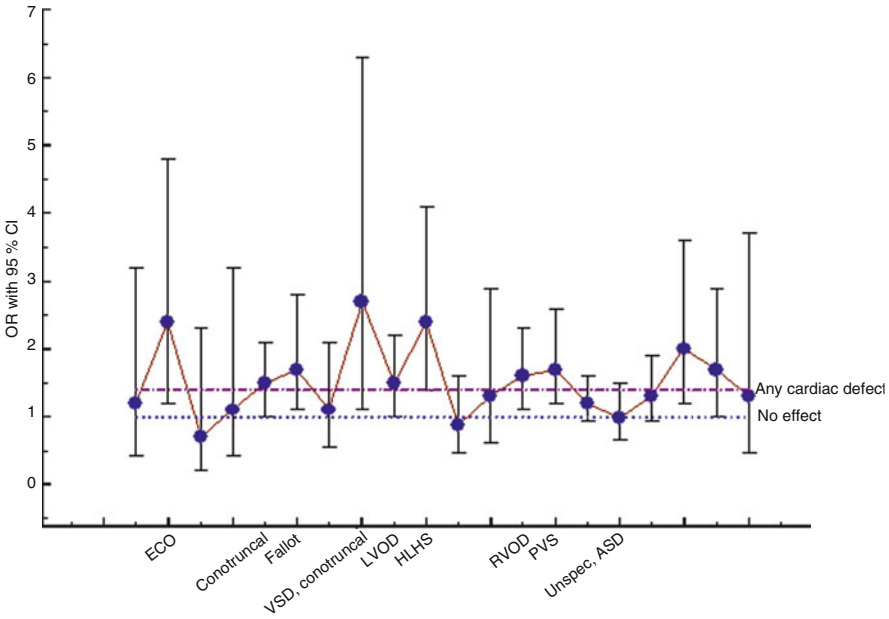


Fig. 4.3 Diagram showing the registered odds ratios (*OR*) with 95 % confidence intervals for 20 different groups of cardiovascular defects according to Broussard et al. (2011). The dashed line gives the average for all these defects and the dotted line the "no effect" line. Cardiovascular types showing significant difference from "no effect" and highlighted in the study are lettered on the X-axis (After Källén and Reis 2016)

(Broussard et al. 2011). These authors worked with data from the National Birth Defects Prevention Study and identified an increased risk for some malformations, including some cardiovascular defects. Figure 4.3 summarizes the odds ratios found for some of the 20 specific cardiac defects studied.

The authors concluded that the nine cardiac types marked were increased in rate after maternal use of opioids. An alternative explanation is that the effect on any cardiac defect of a 40 % increase is due to recall or nonparticipation bias and that the 20 specific types scatter randomly around this value. For one condition (HLHS) the lower confidence limit touches the line for average effect – this is to be expected when 20 comparisons is made.

Various methods have been used to reduce the recall bias phenomenon. One has been to use "sick controls," that is, to make comparisons with another group of malformed infants, e.g., infants with chromosome anomalies or clearly genetic conditions. Parents to such "controls" may, however, have had adequate explanations to the abnormalities and may underreport drug use.

A similar method is to compare different malformations to look for specific associations with drug use. This will probably reduce recall bias but may not eliminate it completely. To be effective the "control" malformations should be of a roughly similar degree of severity as the "case" malformations. Such a technique was early used by Safra and Oakley (1975) in a study of benzodiazepines and orofacial clefts and has been used repeatedly in MADRE (Robert et al. 1994) or SAFE-Med studies (Clementi et al. 2010) from the International Clearinghouse for Birth Defects Surveillance and Research. In the latter studies, data from various congenital malformation registers have been used where exposure information usually had been obtained shortly after the birth of the infants. Also in studies from EUROCAT, a selected malformation has been compared with other malformations with respect to drug exposure, e.g., lamotrigine exposure in infants with orofacial clefts versus infants with other malformations (Dolk et al. 2008).

4.3 Cohort Studies

A second classical epidemiological method is the cohort study (Fig. 4.4). This is based on a group of women with the same exposure, in this case drug use. The rate of outcomes (e.g., malformations in their infants) is compared with the rate of outcomes in a non-exposed cohort. The control cohort is thus a sample of all non-exposed women and will give an idea of the outcome rate (e.g., malformations) among infants of women who were not using the drug in question.

The first problem is usually to identify large enough numbers of exposed women if not information is available for all women in the population (and then a sampling is not needed). We will come back to this problem in Chap. 7. This study design often gives rather small studies with a low power to detect anything but strong effects. A second problem is to identify the outcome under study which sometimes is made by questionnaires or interviews, sometimes from medical documents or registers, e.g., registers of congenital malformations.

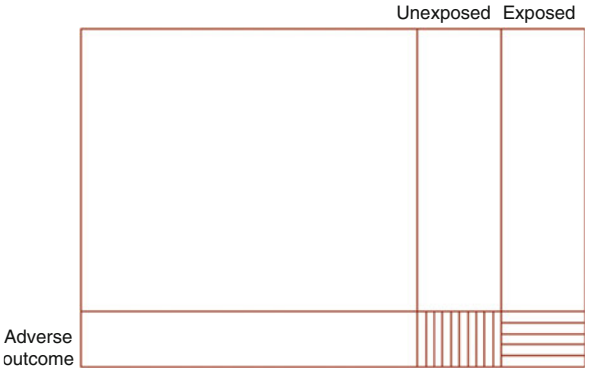


Fig. 4.4 Diagram illustrating the principle of a cohort study. The sizes of the two hatched areas are compared

4.4 Comparison Between Case-Control and Cohort Studies

It is often stated that case-control studies have a higher power to detect associations than cohort studies. This has nothing to do with the type of study but depends on the numbers involved. Generally, the smallest number belongs to the group of exposed outcomes which will therefore contribute most strongly to the uncertainty of the risk estimate. The differences in power between case-control and cohort studies are caused by the difficulties to collect data on large numbers in the latter situation, not on the type of study.

There is a more important difference between case-control and cohort studies. In case-control studies the outcome is decided and it is possible to study many different exposures, e.g., maternal use of different drugs at a certain outcome, e.g., neural tube defects. In cohort studies, the exposure is decided and many different outcomes can be studied, e.g., different malformation types. If the study is restricted to one exposure and one outcome (e.g., use of valproic acid and spina bifida), obviously the two techniques will be equivalent and both are based on sampling from the population in order to get estimates of the expected number of malformed infants after exposure, either estimated from exposure rate in all infants (case-control) or as outcome rate in all infants (cohort). Similarly, expected numbers for the other three groups (unexposed with outcome, exposed, and unexposed without outcome) can be calculated and from these four observed and expected numbers, a chi-square analysis can be made to look for the statistical significance of possible differences in exposure rates between outcome groups or outcome rates between exposure groups (which will be the same). We will come back in greater detail to the evaluation of statistical significances.

4.5 Nested Case-Control Studies

This is illustrated in Fig. 4.5. A crude cohort is first identified where it is likely that the relevant exposure exists. Within that cohort, a case-control study is then performed. Ideally, about half of the crude cohort should be exposed for the factor of interest. This method reduces the number of questionnaires or interviews needed to determine actual exposure – it can be looked upon as a case-control study where exposure rate has been increased by the selection of the crude cohort. Furthermore, the members of the crude cohort may also share common characteristics which otherwise could confound the analysis.

Examples of this type of studies are investigations of infants born after maternal epilepsy, often identified from hospitals specialized in the treatment of epilepsy. With this approach, differential effects of different anticonvulsants can be studied, but no information on the rate in a non-epileptic group of women will exist. In principle this is the background for various registers of anticonvulsants (Russell et al. 2004; Vajda et al. 2004, 2010; Holmes and Wyszynski 2004; Tomson et al. 2004). The crude cohort consists of women with epilepsy and then within that group, risks after exposure to a specific anticonvulsant compared with other anticonvulsants or no anticonvulsants can be made.

4.6 The Sibling Approach

The background to these studies is to compare outcomes in two pregnancies of the same mother, one exposed and the other non-exposed. In this way, the effect of fixed characteristics of the woman can be removed, notably of her genetics. On the other hand, disease status, smoking, body mass index (BMI), and many other things can

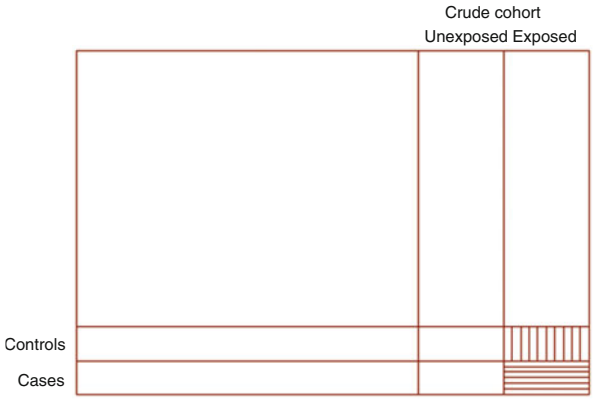


Fig. 4.5 Diagram illustrating the principle of a nested case-control study. Within a crude cohort containing both exposed and non-exposed individuals, a case-control study is performed, comparing the sizes of the hatched areas

have changed between the two pregnancies. Another drawback is that at least two pregnancies are needed for the study and the results may not be applicable for women with only one pregnancy. A great problem is that one not only has to ascertain the drug use in one of the pregnancies but also verify that no drug was used in the other pregnancy. It is definitely a valuable methodology, but one has to be careful in the interpretation of the results.

Most sibling studies on drug use are based on prescription registers. A recent example is the study by Furu et al. (2015) on SSRI drugs and venlafaxine. In this Nordic study, one found an increased risk of a congenital malformation after drug exposure which, however, disappeared in the sibling study – the estimate for any congenital malformation was 1.17 (1.05–1.26) in the covariate-adjusted analysis and 0.92 (0.72–1.17) in the sibling-controlled analysis. Among the 36,772 infants exposed to the drugs in question, only 980 entered the sibling study.

As will be discussed in greater detail later in this book, all information on drug use may have two shortcomings. One is that a woman may have used a drug during pregnancy without this being identified. In interview or questionnaire studies, she may not have told about the drug use and in prescription studies she may have had access to and used drugs which were prescribed much earlier; most drugs have a shelf life of many years. The second problem is that she may not have used a drug which we think she has used. It is rather unlikely that she did not use a drug which she says that she used, but she may have mistaken the time when she used it, notably if data are collected months after delivery. In prescription studies it is a definite risk that she bought the drug but did not use it, especially not during early pregnancy.

Some unpublished data on the effect of antidepressants on preterm birth in singletons will be presented. They are based on data from the Swedish Medical Birth Register for 2005–2013. There were 9595 singleton pregnancies where the mother had reported the use of antidepressants in early pregnancy. Among them 2786 had siblings during the study period; the total number of unexposed siblings was 2922 where the mother had stated the use of any other drug than an antidepressant, including vitamins. The adjusted odds ratio for preterm delivery among the total group was 1.53 (95 % CI 1.42–1.66) while among the antidepressant-exposed infants with siblings, the odds ratio was 1.34 (95 % CI 1.16–1.56). This indicates that the group with siblings is a selected subgroup. The unexposed siblings did not differ from the population: odds ratio = 1.06 (95 % CI 0.93–1.21). Sibling studies may thus give biased data. In this example, however, the study indicates that the effect on preterm birth is at least partly drug induced.

References

Botto LD, Krikov S, Carmichael SL, Munger RG, Shaw GM, Feldkamp ML, National Birth Defects Prevention Study (2015) Lower rate of selected congenital heart defects with better maternal diet quality; a population-based study. Arch Dis Child Fetal Neonatal Ed 101:F43–F49. doi:10.1136/archdischild-2014-308013

Broussard CS, Rasmussen SA, Reefhuis J, Friedman JM, Jann MW, Riehle-Colarusso T, Honein MA, National Birth Defects Prevention Study (2011) Maternal treatment with opioid analgesics and risk of birth defects. Am J Obst Gynecol 204:314.e1–e11

Clementi M, Di Gianantonio E, Cassina M, Leoncini E, Botto LD, Mastroiacovo P, Safe-Med Study Group (2010) Treatment of hyperthyroidism in pregnancy and birth defects. J Clin Endocrinol Metab 95:E337–E341

Dolk H, Jentink J, Loane M, Morris J, de Jong/van den Berg LT, EUROCAT Antiepileptic Drug Working Group (2008) Does lamotrigine use in pregnancy increase orofacial cleft risk relative to other malformations? Neurology 71:714–722

Furu K, Kieler H, Haglund B, Engeland A, Selmer R, Stephanson O et al (2015) Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population-based cohort study and sibling design. BMJ 350:h1798. doi:10.1136/bmj.h1798

Holmes LB, Wyszynski DF (2004) North American antiepileptic drug pregnancy register. Epilepsia 45:1465

Källén B, Reis M (2016) Ongoing pharmacological management of chronic pain in pregnancy. Drugs 76:915–924

Robert E, Vollset SE, Botto L, Lancaster PA, Merlob P, Mastroiavo P et al (1994) Malformation surveillance and maternal drug exposure. Int J Risk Saf Med 6:75–118

Russell AJC, Craig JJ, Morrisson P, Irwin B, Waddell R, Parsons L et al (2004) UK epilepsy and pregnancy group. Epilepsia 45:1467

Safra JM, Oakley GP Jr (1975) Association between cleft lip with or without cleft palate and prenatal exposure to diazepam. Lancet ii:478–480

Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Perucca E, Collaborative EURAP Study Group et al (2004) EURAP: an International registry of antiepileptic drugs and pregnancy. Epilepsia 45:1463–1464

Vajda F, Lander C, O'Brien T, Hitchcock A, Graham J, Solöinas C et al (2004) Australian pregnancy register of women taking antiepileptic drugs. Epilepsia 45:1466

Vajda FJ, Graham JE, Hitchcock AA, O'Brien TJ, Lander CM, Eadie MJ (2010) Is lamotrigine a significant human teratogen? Observations from the Australian pregnancy register. Seizure 19:558–561

Yau WP, Mitchell AA, Lin KJ, Werler MM, Hernández-Dias S (2013) Use of decongestants during pregnancy and the risk of birth defects. Am J Epidemiol 178:198–208

Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SJ et al (2001) The national birth defects prevention study. Public Health Rep 116(Suppl):32–40

Pregnancy Outcomes with the Exception of Congenital Malformations

5

Most interest has been paid to the risk of congenital malformations (see Chap. 6). Maternal use of drugs during pregnancy may, however, affect pregnancy outcome in many ways. We will discuss the most important such outcomes and methods to identify them in epidemiological studies.

5.1 Maternal Pregnancy Complications

Such complications are usually identified from medical records or – if available – from medical birth registers and are usually defined by International Classification of Diseases (ICD) codes. Interview or questionnaire data are less reliable and may be biased.

5.1.1 Preeclampsia

Mild to severe preeclampsia is registered in about 4 % of all pregnancies in the Swedish Medical Birth Register. Among women who had used antidepressants in early pregnancy, 5.4 % had preeclampsia and among women who did not use antidepressants, 3.9 % did – a crude odds ratio is 1.40 (95 % CI 1.32–1.50). There are many factors which can influence both the use of antidepressants in early pregnancy and preeclampsia later in pregnancy. With increasing maternal age, antidepressant use increases and so does the risk for preeclampsia. The use of antidepressants and also the preeclampsia risk decreases with parity. The same trend of an increase with maternal prepregnancy BMI is seen for both phenomena. Maternal smoking is associated with the use of antidepressants but has a protective effect on the development of preeclampsia. After adjustment for these factors and year of delivery, the odds ratio for preeclampsia in women who had used antidepressants in early pregnancy was 1.26 (95 % CI 1.12–1.34), and these factors thus only explained one-third of the crude risk. A strong risk factor for preeclampsia is

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_5

chronic hypertension which is evident from the odds ratio for preeclampsia when the woman used antihypertensives in early pregnancy (4.22, 95 % CI 3.86–4.21). An existing association between chronic hypertension and use of antidepressants could explain the effect of the latter, but exclusion of women who had used antihypertensives changed the odds ratio only little (1.24, 95 % CI 1.18–1.33). The link between the use of antidepressants and preeclampsia seems not to be due to chronic hypertension. If the effect is caused by the drugs, the underlying disease or some common factor is not clear. Palmsten et al. (2013) claimed that untreated depression during pregnancy was associated with an increased risk for preeclampsia which was further increased if antidepressant drugs (notably TCA or SNRI) had been used.

Preeclampsia is a serious pregnancy complication which often results in preterm birth and increases the risk for post-pregnancy hypertension, cardiovascular disease, and chronic kidney disease (Vest and Cho 2014). It can also develop into eclampsia even if this is relatively rare in developed countries or into the life-threatening HELLP syndrome.

5.1.2 Placenta Previa

Placenta previa is a less common complication than preeclampsia and is due to an abnormally low implantation of the embryo so the placenta will lie below the fetus in the uterus. In the Swedish Medical Birth Register, this diagnosis was given to 0.3 % of all women who did not use antidepressants and 0.5 % of all women who used antidepressants. The crude odds ratio for placenta previa after maternal use of antidepressants is 1.42 (95 % CI 1.15–1.76). The risk of placenta previa increases with maternal age or maternal smoking but is not much influenced by parity or BMI. Adjustment for year of delivery, maternal age, parity, smoking, and BMI changes the odds ratio only little: 1.46 (95 % CI 1.17–1.81). The condition usually results in a cesarean section (96 % of the diagnosed cases in the Swedish Medical Birth Register).

5.1.3 Placenta Abruption

This condition means that the placenta is detached, partially or completely, before the baby has been born. It results in major bleeding and threatens the life of the baby. In the Swedish Medical Birth Register, this complication was registered in 0.3 % among women who did not use antidepressants and 0.4 % in women who used such drugs. The crude odds ratio was 1.22 (95 % CI 0.97–1.54), thus not statistically significant. Also the risk of this complication increases with maternal age, is higher at first parity than at second to third parity, increases with smoking, and is increased at low BMI. After adjustment for these factors, the odds ratio decreased to 1.04 (95 % CI 0.80–1.35), so the suggested effect seemed mainly to be due to maternal characteristics (Reis and Källén 2010).

5.1.4 Other Pregnancy Complications

Among other pregnancy complications which may be related to drug use during pregnancy (or underlying disease) can be mentioned hyperemesis gravidarum, gestational diabetes, premature rupture of the membranes, and bleeding around delivery. The possibly two-way association between maternal complication and drug use can be exemplified with hyperemesis. This condition is treated with various drugs, including antihistamines, but there are data published which suggests that the use of antihistamines could increase the risk for the adverse effects of hyperemesis (Fejzo et al. 2013).

Many of these complications are more likely to be affected by maternal drug use in middle and late pregnancy than in early pregnancy, but some early effects are possible, for instance, by affecting placenta development. Many pregnancy complications result in an increased use of instrumental deliveries, notably cesarean sections.

5.2 Spontaneous Abortion

Spontaneous abortion or miscarriage is the death of an embryo or fetus before it has reached the age of becoming an infant. This age varies somewhat between populations. The classical limit is drawn when an infant born has a chance to survive after birth. Today the limit is around 22 weeks and this limit is therefore used for instance in Sweden as the lower gestational age limit to define an infant which was born dead. Historically, the limit was higher. Up to 2002 an age limit of 28 weeks was used in Sweden to define a dead fetus as an infant – if it was alive at birth it was defined as an infant, also if born before the 28th week. The variability in definition will not much affect rate estimates as most stillbirths occur late in pregnancy and most spontaneous abortions occur early.

One can also discuss if there should be a lower age limit in analyses of miscarriages. Most likely, a high percentage of fertilized eggs and early embryos – which thus have the potential to develop into an infant – stop developing and the pregnancy may never be realized by the woman. It is possible to identify very early pregnancies with biochemical methods, and some studies have indicated that nearly half of them will never reach the age limit for an infant; most will succumb very early before the woman knows she is pregnant.

In most circumstances studies are made only on miscarriages after the time point when the woman knows about her pregnancy. This will in itself introduce a degree of uncertainty because of the variation in that time point between women. Information on miscarriages is usually obtained from interviews or questionnaires; in some countries it is possible to use medical diagnosis registers which will only identify those women with miscarriage who have searched medical advice or have been hospitalized. When different groups of women are compared (for instance, with reference to drug use), differences in these features may exist which make comparisons uncertain.

From an epidemiological point of view the relatively high number of miscarriages is an advantage. At least 10–15 % of all known pregnancies ends with a miscarriage. This relatively high rate is balanced by the difficulty to identify miscarriages without bias and, as we will see, there is a problem in the analyses of the data.

Most miscarriages occur early in pregnancy and the majority of these represent abnormal embryos, often with gross chromosome anomalies. Later occurring miscarriages (after week 12–14) are usually normal fetuses that miscarry because of some maternal condition. They are actually biologically more closely related to stillbirths than to early miscarriages.

If we suppose that we have succeeded to adequately identify all miscarriages in two groups of women and want to compare the miscarriage rates in the groups, we have to consider what type of denominator should be used. Strictly speaking, the risk for a woman to miscarry is the number of miscarriages occurring among the number of pregnancies at risk at that time of the pregnancy. Pregnancies at risk include pregnancies which will go to delivery but also pregnancies which will, at a later stage of pregnancy, be interrupted by an induced abortion (legal or in some populations illegal). Therefore, exposure rates among pregnancies that miscarry should be compared with the exposure rate among all intrauterine pregnancies which were alive at the time of the miscarriage. This will result in the need for a type of life-table analysis. If the exposure rate among miscarriages is only compared with the exposure rate in pregnancies which continue to delivery, the result will be biased if induced abortion is associated with the exposure under study (as may notably be the case with many psychoactive drugs or drugs with a suspected teratogenic activity). A further discussion of this problem can be found in Källén (2012).

It is often difficult to perform a life-table analysis of miscarriage rate due to lack of adequate data. Various shortcuts have been suggested, e.g., to use as denominator the sum of spontaneous and half of the induced abortions plus births – this means that one supposes that as an average half of induced abortions occur after the gestational age of the studied spontaneous abortions (Susser 1983).

This problem can be illustrated with data from an old study, based on prospectively collected information on drug use during pregnancy (Kullander and Källén 1976). The use of psychoactive drugs was studied. The use of such drugs was nearly twice as common among women with an unwanted pregnancy as among women with a wanted pregnancy. The drug use rate among women who miscarried was 10 %, among women who gave birth only 6 %, but among women who later had an induced abortion it was about 30 %. Estimates indicated that the nearly doubling of the exposure rate in women who will miscarry was due to the effect of the association between drug use and induced abortions. Similar results were reached for maternal smoking and early – but not late – miscarriages.

The literature on maternal drug use and miscarriage risk has usually not taken these complications into consideration. It is typical that in a review and meta-analysis of the problem of antidepressant use during pregnancy and miscarriages, no discussion was made of this basic problem (Hemels et al. 2005). Other studies excluded women with an induced abortion (e.g., Nakhai-Pour et al. 2010) which of course does not solve the problem.

A moderate increase of the miscarriage risk is difficult to demonstrate with certainty because of these complications.

5.3 Stillbirth and Infant Death

Biologically, stillbirths and late spontaneous abortions represent a continuum and the borderline is mainly of administrative nature. Intrauterine death may be the result of some of the pregnancy complications mentioned above, e.g., placenta abruption, but often the mechanism is not known and a specification of the cause of death is often not made. The rate of stillbirths depends somewhat on the definition of the lower age limit, but most stillbirths occur much later. In developed countries with a well-functioning prenatal care, the stillbirth rate is 0.3–0.4 %. The risk for intrauterine death increases with maternal age, is higher at first parity compared with higher parities, and increases with smoking and also with maternal overweight or obesity. Relatively few drugs have been linked to intrauterine deaths, for example, ACE inhibitors and angiotensin II antagonists used as antihypertensives (Pucci et al. 2015).

Among infants born of women using antidepressants, 0.40 % was stillborn and among women not using such drugs, 0.35 %. The crude odds ratio was 1.12 (95 % CI 0.90–1.42) and after adjustment for year of delivery, maternal age, parity, smoking, and BMI, it decreased to 1.04 (95 % CI 0.81–1.33). In spite of the rather large size of the study (3420 stillbirths among which 78 were born of women using antidepressants), the upper confidence limit permits a 33 % excess risk even though the risk estimate is rather close to 1.0.

The distinction between stillbirths and death immediately after birth is not always clear. Sometimes, one uses the concept of perinatal death, including both stillbirths and early neonatal deaths (<7 days after birth). Such deaths are usually identifiable from obstetric records, but if survival should be followed for a longer time, linkage with registers of death gives more complete results. Neonatal death risk is strongly associated with very short gestational duration and gross congenital malformations.

5.4 Gestational Duration and Birth Weight

Gestational duration and birth weight are standard information on all infants born in most data sources and invite analyses.

Classically, gestational duration is calculated from the last menstrual period (LMP) as stated by the woman. Nowadays, gestational duration is often estimated by fetal size determined with sonography. This gives a better estimate even though some minor uncertainties may exist. After in vitro fertilization, exact gestational age is known from the date of embryo transfer and length of embryo incubation, and this could be compared with the age estimated from sonography (Källén et al. 2013a). Even though in general the estimates agreed well, it could be shown that overweight

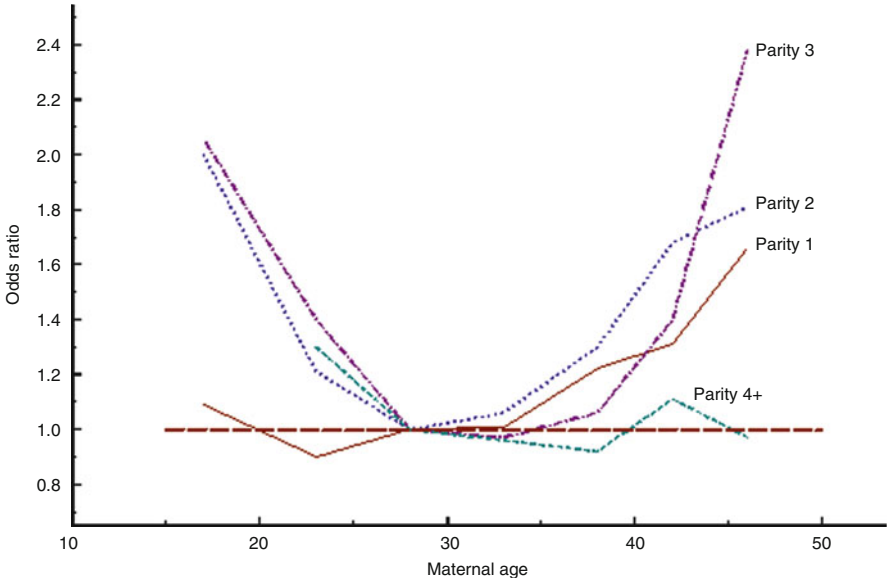


Fig. 5.1 Diagram showing the relationship between maternal age and parity and the risk for preterm birth, adjusted for year of birth, smoking, and body mass index

or obesity could influence the exactness of the estimate. Infants born growth retarded may also have been sensitive for estimate errors.

Gestational duration is shorter in twins or higher-order births than in singleton births. For this reason analyses of effects on gestational duration are usually restricted to singleton births.

A common measure consists of preterm birth, that is, births less than 37 weeks. A clinically more important distinction is very preterm birth, shorter than 32 weeks, which of course is a more rare (less than 1 %) but for the neonate a more critical condition. The rate of preterm births varies markedly between different populations, partly because of socioeconomic conditions, partly because of the quality of prenatal and delivery care. In the Swedish Medical Birth Register, 5.2 % of male infants and 4.7 % of the female infants were preterm , and the total preterm rate was 4.9 %. Some extremely preterm infants are, however, missing from the register.

Another method is to determine the mean gestational duration. We can illustrate this with a comparison of the two measurements in singleton infants born by women who had or had not used antidepressants during pregnancy. After antidepressants use, the rate of preterm births increased from 4.9 to 7.8 %, and at the same time the mean gestational duration decreased from 39.4 to 39.0 weeks, a difference of about 3 days. The difference in the rates of preterm birth is clinically important; the difference in mean gestational duration is hardly that.

Numerous factors affect gestational duration and rate of preterm births. In Fig. 5.1 it can be seen that the relationship between maternal age and preterm birth varies with parity and is nonlinear and of different shape at different parities.

Table 5.1 Relation between preterm birth and maternal smoking status

	<32 weeks		<37 weeks	
	OR	95 % CI	OR	95 % CI
No smoking	1.00	Reference	1.00	Reference
Smoking <10 cigarettes/day	1.62	1.48–1.78	1.11	1.07–1.48
Smoking ≥10 cigarettes/day	2.00	1.80–2.42	1.79	1.68–1.41

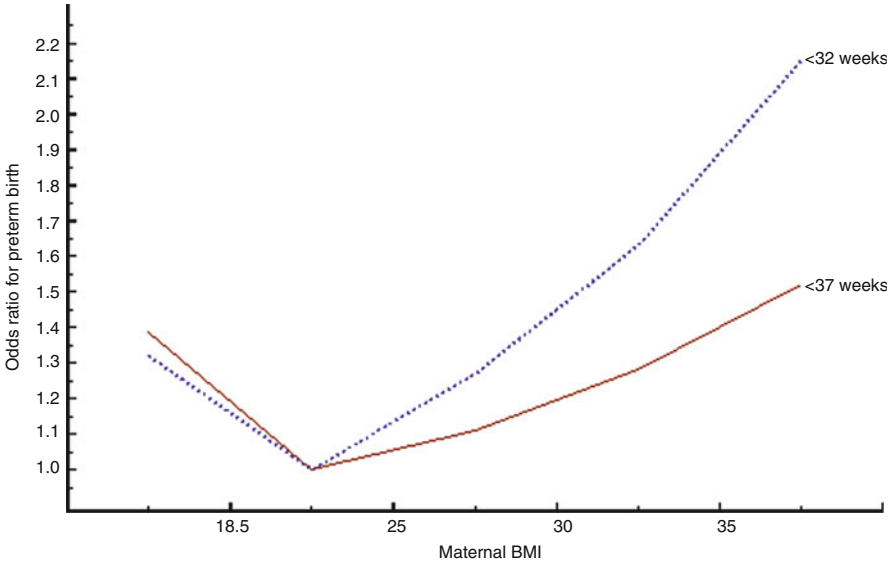


Fig. 5.2 Risk of preterm birth (<32 weeks and <37 weeks) according to maternal BMI, adjusted for year of delivery, maternal age, parity, and smoking

The effect of maternal smoking (in early pregnancy) on preterm birth is dose dependent and stronger for <32 weeks than for <37 weeks as is seen in Table 5.1.

Also maternal BMI has an effect on the preterm risk as seen in Fig. 5.2. There is a moderate excess of preterm births when the mother has a low BMI, and with a higher than normal BMI (18.5–24.9) the risk increases and is higher for preterm births <32 weeks than for <37 weeks.

Smoking and overweight or obesity are related to the socioeconomic situation. In Sweden the socioeconomic differences are moderate, and pregnancy care is the same for everyone and is free of charge. There are other factors associated with socioeconomic level which may play a role, e.g., nutrition. In a recent study (Smith et al. 2015), it was shown that a diet rich in fruit and vegetables or with Mediterranean characteristics reduced the preterm rate (32–36 weeks). In countries with large social differences, socioeconomy probably plays a larger role than in Sweden. In the paper by Smith et al., no effect of alcohol or recreational drugs was seen. In abuse situations, such effects are seen, however. Thus preterm birth is a feature of the fetal alcohol syndrome, and abuse of opioids also increases

the risk for preterm birth together with a number of other ill effects on the neonate.

Birth weight is a consequence of two factors: gestational age at birth and intrauterine growth. An advantage in using birth weight as an outcome in an epidemiological study is the relative certainty of the information. The usual definition of very low birth weight is <1500 g and of low birth weight, <2500 g. This variable is usually studied in singleton births because twins or higher-order births have lower birth weight in each gestational week than singleton infants have.

5.5 Intrauterine Growth

Intrauterine growth can be followed with sonography, but the end result of a disturbance of intrauterine growth is shown by the infant weight given the week in which it was born. In order to evaluate this, normal growth diagrams are used. These are specific for the studied populations. The definition of intrauterine growth retardation also varies between studies. Often the tenth percentile is used to define growth retardation, sometimes two standard deviations from the mean weight. Usually such growth diagrams are based on the mean and distribution of the birth weights for each gestational week. This means that pathological conditions are included in what is supposed to be a normal population. Other methods have instead used weight modes for each week and estimated the standard deviation as 12 % of that weight – the 12 % come from 40-week pregnancies where the vast majority is normal (Källén 1995). One method was based on intrauterine sonographic weight estimates of infants which were later born at term and were normal (Marsál et al. 1996).

As long as the same growth curves are used for all studied groups, their exact appearance is of less importance. There is, however, a principal difference between the different manners to construct the curves. In most curves based on data in the newborns including the Källén curve (which can be standardized for infant sex and maternal parity), one compares the birth weight of a certain infant with the most common birth weight or the mean weight at that gestational age. As infants born preterm are often growth retarded, this factor is eliminated and the estimate tells if an infant is more growth retarded than the majority of infants born that week. The two growth curves of the Källén and the Marsál diagrams will be identical in term infants, but the latter will lie above the former for preterm births as a result of the fact that infants born preterm are often growth retarded.

Whichever growth curve is used, one can identify three groups of infants: small for gestational (SGA), appropriate for gestational age (AGA), and large for gestational age (LGA) infants. Even though there is an association between SGA and preterm births and LGA and post-term births, it is useful to be able to study the two phenomena independently: does a drug affect the gestational age at which the infant is born or does it affect intrauterine growth without necessarily affecting gestational duration?

Disturbed intrauterine growth resulting in small-for-date babies has important implications for the later development of the child.

Table 5.2 Effect of maternal antidepressant use on respiratory diagnoses in the infant according to pregnancy duration

Group of infants	OR	95 % CI
All	1.71	1.62–1.81
Post-term infants (≥42 weeks)	1.20	0.89–2.62
Term infants (37–41 weeks)	1.89	1.77–2.03
Preterm infants (<37 weeks)	1.07	0.95–1.22
Very preterm infants (<32 weeks)	1.28	0.82–2.02

5.6 Other Body Dimensions

Like body weight, infant length and head circumference depend on gestational duration. These variables are seldom used in epidemiological studies of the effect of maternal drug use. The effect of maternal anticonvulsant use on infant head circumference was studied by Almgren et al. (2009). They adjusted head circumference for 500 g birth weight classes and calculated the deviation for each exposed infant expressed as standard deviations from the mean head circumference in the birth weight class to which the infant belonged. In this way it was possible to compare infants exposed for anticonvulsants with unexposed infants and also to compare the effect of different anticonvulsants. The clinical significance of a moderately reduced head circumference is uncertain.

5.7 Perinatal Morbidity

Maternal drug use may be associated with perinatal morbidity. Such effects are known for some drugs of abuse, e.g., heroin, but have also been described for instance after maternal use of antidepressants. They may appear as intrauterine asphyxia, respiratory problems in the neonate, hypoglycemia, jaundice, symptoms of central nervous system disturbances, etc. and are identified from diagnoses in medical records or from medical birth registers. Interview or questionnaire data are less reliable and may be biased. A summary evaluation of the status of the infant the first few minutes after birth is given by the Apgar score (0–10), which is usually given after at least 1 and 5 min after birth. Often <7 at the 5-min test is used as a sign of low Apgar score, but in some studies the 1-min score were used. The predictive capacity of the Apgar score has been debated but is usually regarded as of value (Casey et al. 2001; Stuart et al. 2011; Tweed et al. 2015).

An increased risk for perinatal morbidity may be a consequence of an increase in preterm births or a direct effect on the fetus. Among all infants 2.9 % had one or more such diagnoses, 4.6 % among post-term infants, 2.6 % among term infants, 26.5 % among infants born <37 weeks, and 67.1 % among infants born <32 weeks. Table 5.2 shows the effect of maternal use of antidepressants on respiratory diagnoses in singleton newborns. It can be seen that the strongest effect is seen on term infants while no certain effect is seen on post-term or preterm infants. This speaks

for a direct effect of the drug which is partly hidden by the increased risks associated with post-term or preterm births.

Some neonatal diagnoses are uncommon and large materials are needed to study them. An example is PPHN which in pregnancies with more than 34 weeks duration occurs in only some three per 1000 births. Another example is necrotizing enterocolitis which is also a rare complication, typically a consequence of very preterm birth.

5.8 Long-Term Effects

It has long been known that abuse of alcohol or recreational drugs may cause long-term effects on the cognitive development of the exposed child and also other developmental disturbances. Less is known about such effects of medically used drugs but has been described for anticonvulsants (notably valproic acid) and also for some other psychoactive drugs. During the last decade or so, other long-term effects of maternal drug use have been discussed, e.g., effects on the risk of childhood asthma, ADHD, and autism.

Different approaches can be used in order to identify long-term outcomes. In small studies, various psychological tests can be applied in order to identify effects of maternal drug use. Usually mean values are compared between exposed and unexposed groups, a method which is un-sensitive if the exposure has caused an increased risk of an uncommon outcome. The mean intelligent quotient may be nearly the same in the two study groups, but the exposure may have increased the risk for mental retardation. This is a situation similar to that when mean gestational duration is compared with the rate of preterm birth (see above).

With modern health and other registers, it is sometimes possible to follow a large number of children which have been exposed in utero for specific drugs. This makes it possible to identify also risk increases for less common outcomes. For this purpose, various sources can be used. We can exemplify this with childhood asthma. Some studies have used hospital discharge registers to identify such disease, sometimes also diagnoses from outpatient clinics. Other studies have used registers of filled prescriptions and regarded at least repeated filling of prescriptions for anti-asthmatics as evidence for asthma in the child (Källén et al. 2013b). In selected populations (often from selected strata of the population), data from medical insurance systems have been used. There are weaknesses with all methods. Hospital discharge diagnoses have a tendency to identify severe cases which have needed hospitalization. Outpatient information may be biased if maternal characteristics affect the probability that a child with perhaps light asthma is taken to medical care. This will also affect the results of studies based on prescription registers where the problem also exists that anti-asthmatics may have been prescribed for other conditions than asthma.

In studies of mental retardation or other severe neuropsychiatric conditions, health registers have been much used, for instance the Danish psychiatric register (Sørensen et al. 2013). Prescription registers can be used to identify children with

conditions when specific drugs are used for the conditions. An example is the use of methylphenidate or similar drugs at ADHD (Källén et al. 2013c). In some populations it is possible to study school results, e.g., at the end of compulsory school, based on national registers of school marks (Forsberg et al. 2011).

Other health registers may exist which make it possible to study specific long-term outcomes. One example is cancer registers which give good information on the occurrence of malignant tumors.

In order to follow the children through life, valid identification is necessary. Notably in the Scandinavian countries, this is easily done with the help of the personal identification number every person living in the country gets at or immediately after birth. Such numbers are widely used in society and in all health care. In order to protect patient privacy, the identification number can be changed to other, neutral numbers, but that has to be done in a similar way in all health registers to be useful. In many populations probability linkage between registers has to be made which is more complex and uncertain than the use of individual identification numbers.

All studies of prenatal effects of drug use on long-time development of the child have complications. A genetic component in the disease for which the drug was taken can transfer susceptibility for the disease to the child. The situation during the child's early life may also be affected by the maternal disease and cause disturbances in the child's development.

We can exemplify this problem with studies on maternal use of antibiotics during pregnancy and childhood asthma. Many studies have shown that such an association exists, but as maternal asthma is associated with antibiotic use and a genetic component for asthma exists, the association could be spurious. This was recently shown in a large study by Örtqvist et al. (2014) where a sibling analysis seemed to remove the association completely. In this case, about one-third of the material was eligible for sibling analyses.

Permanent life-long impairment of the child's function is perhaps of greater importance than many congenital malformations which can easily be corrected.

References

Almgren M, Källén B, Lavebratt C (2009) Population based study of antiepileptic drug exposure in utero – influence on head circumference in the newborn. Seizure 18:672–675

Casey BM, McIntire DD, Leveno KJ (2001) The continuing value of the Apgar score for the assessment of newborn infants. N Engl J Med 344:467–471

Fejzo MS, Magtira A, Schoenberg FP, MacGibbon K, Mullin P, Romero R et al (2013) Antihistamines and other prognostic factors for adverse outcome in hyperemesis gravidarum. Eur J Obst Gynecol Reprod Biol 170:71–76

Forsberg L, Wide K, Källén B (2011) School performance at age 16 in children exposed to antiepileptic drugs in utero – a population-based study. Epilepsia 52:364–369

Hemels ME, Einarson A, Koren G, Lanctot G, Einarson TR (2005) Antidepressant use during pregnancy and the rates of spontaneous abortions: a meta-analysis. Ann Pharmacother 39:803–809

Källén B (1995) A birth weight for gestational age standard based on data in the Swedish Medical Birth Register. Eur J Epidemiol 11:601–605

Källén B (2012) Chapter 26. Human studies – epidemiologic techniques in developmental and reproductive toxicology. In: Hood R (ed) Developmental and reproductive toxicology. A practical approach, 3rd edn., Informa Healthcare, London, pp 645–647

Källén B, Finnström O, Nygren K-G, Otterblad Olausson P (2013a) Maternal and fetal factors which affect fetometry: use of in vitro fertilization and birth register data. Europ J Obstet Gynecol Reprod Biol 170:372–375

Källén B, Finnström O, Nygren K-G, Otterblad Olausson P (2013b) Maternal drug use during pregnancy and asthma risk among children. Paed Allergy Immunol 24:28–32

Källén B, Borg N, Reiss M (2013c) The use of central nervous system active drugs during pregnancy. Pharmaceuticals 6:1221–1286

Kullander S, Källén B (1976) A prospective study on drugs and pregnancy. I Psychopharmaca Acta Obstet Gynecol Scand 55:25–33

Marsál K, Persson PH, Larsen T, Lilja H, Selbing A, Sultan B (1996) Intrauterine growth curves based on ultrasonically estimated foetal weights. Acta Paediatr 85:843–848

Nakhai-Pour HR, Broy P, Bérard A (2010) Use of antidepressants during pregnancy and the risk of spontaneous abortion. CMAJ 182:1031–1037

Örtqvist AK, Lundholm C, Kieler H, Ludvigsson JF, Fall T, Ye W et al (2014) Antibiotics in fetal and early life and subsequent childhood asthma: nationwide population based study with sibling analysis. BMJ 349:g6979. doi:10.1136/bmj.g6979

Palmsten K, Huiybrechts KF, Michels KB, Williams PL, Mogun H, Setogushi S et al (2013) Antidepressant use and risk for preeclampsia. Epidemiology 24:682–691

Pucci M, Sarween N, Knox ED, Lipkin G, Martin U (2015) Angotensin-converting enzyme inhibitors and angiotensin//receptor blockers in women of childbearing age: risk versus benefits. Expert Rev Clin Parmacol 8:221–231

Reis M, Källen B (2010) Delivery outcome after maternal use of antidepressant drugs in pregnancy: an update using Swedish data. Psychol Med 11:1723–1733

Smith LK, Draper ES, Evans TA, Field DI, Johnson SJ, Manktelow BN et al (2015) Associations between late and moderately late preterm birth and smoking, alcohol, drug use and diet: a population-based case-cohort study. Arch Dis Child Fetal Neonatol Ed 100:F486–F91. doi:10.1136/archdischild-2014-207265

Sørensen MJ, Grenborg TK, Christensen J, Parner ET, Vestergaard M, Schendel D et al (2013) Antidepressant exposure in pregnancy and risk of autism spectrum disorders. Clin Epidemiol 5:449–459

Stuart A, Otterblad Olausson P, Källén K (2011) Apgar scores at 5 minutes after birth in relation to school performance at 16 years of age. Obstet Gynecol 118:201–208

Susser E (1983) Spontaneous abortion and induced abortion: an adjustment for the presence of induced abortion when estimating the rate of spontaneous abortion from cross-sectional studies. Am J Epidemiol 117:305–308

Tweed EJ, Mackay DF, Nelson SM, Cooper S-A, Pell JP (2015) Five-minute Apgar score and educational outcomes: retrospective cohort study of 751 369 children. Arch Dis Child Feral Neonatal Ed 101:F121–F126. doi:10.1136/archdischild-2015-308483

Vest AR, Cho LS (2014) Hypertension in pregnancy. Curr Atheroscel Rep 16:395. doi:10.1007/s11883-013-0395-8

Congenital Malformations

6

6.1 Definition

There is no generally agreed-upon definition of a congenital malformation. It should involve a structural abnormality which has developed during intrauterine life. Inborn errors of metabolism, nearly always genetically caused, without structural abnormalities should therefore not be included. The same is true for some other birth defects like congenital cerebral palsy or mental retardation without a structural background, but a structural congenital malformation may cause mental retardation. The concept of structural abnormality is not well defined. How much should a "malformed" individual differ from the "normal" individual? Variations for instance in external ear morphology or the presence of a four-finger line in the hand are no congenital malformations but normal variants even though both are overrepresented for instance in individuals with Down syndrome. Nevus is really a skin malformation but practically all of us carry nevi. A restriction to more severe conditions is needed but will be somewhat arbitrary. Can a morphologically patent oval foramen in the heart (which is a very common phenomenon, perhaps in 25 % of all individuals and usually without any clinical significance) be regarded as a malformation? The normal closure has actually not taken place, and the condition can under rare circumstances cause problems and may have to be surgically corrected.

Minor variants should not be included in the concept of congenital malformations, but the definition will be vague. If many variants are present simultaneously, this may indicate a disturbed morphogenesis. We have already mentioned that some variants are more common at Down syndrome than in the population, and there is a list of further such signs. Such phenomena are often called dysmorphology, and characteristic signs have been described after maternal abuse of alcohol (fetal alcohol syndrome, FAS) or after maternal use of some anticonvulsants, first described by Bénthenod and Frédérich (1975), Seip (1976), and Hansson et al. (1976). In order to identify dysmorphology, a qualified examination of the child by a specialist is needed. This is usually relevant only in small studies of specific drugs and should be made "blind" as the evaluation is somewhat subjective.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_6

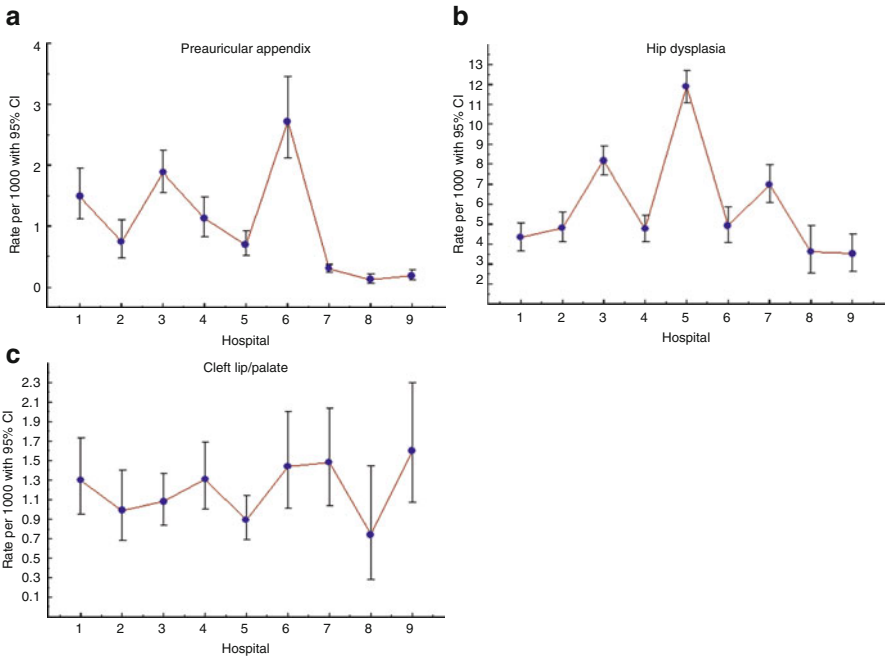


Fig. 6.1 Diagrams showing variation in registered rates of two minor (preauricular appendix (**a**) and hip dislocation (**b**)) and one major (cleft lip/palate (**c**)) malformations in nine teaching hospitals. Hospitals 1–4 are all Stockholm hospitals

Also among conditions which are usually called congenital malformations, there exists a marked variability in severity and some are of relatively modest clinical significance. Often a distinction is made between "major" and "minor" malformations, and analyses are often restricted to major ones. There are two reasons for this. One is that the minor ones make up a substantial part of the total and can hide effects on major malformations; the other is that there is a much higher variability in the reporting of minor malformations than of major malformations. This can be illustrated by data from Sweden where information on congenital malformations is collected from multiple register sources. In Fig. 6.1 the rates of three malformations are compared between nine large hospitals (10,000–70,000 births during the observation period 2005–2013). Two malformations are minor (preauricular appendix and hip dysplasia); one is major (cleft lip or cleft palate). A large variation in rates is seen for the first two conditions while the variability for the third is much less and may be random.

6.2 Major and Minor Malformations

A common definition of a major malformation is that it should be potentially lethal, need surgery or other treatment, or give major cosmetic problems. For many conditions, this definition is quite adequate (e.g., spina bifida, major heart defects, cleft

lip/palate, limb reductions), for others it is more difficult to use the definition. Postaxial polydactyly needs surgery but it is often a very quick and minor operation, and the malformation is therefore sometimes not counted as major.

If the description of the malformation is detailed, it is usually possible to classify it as major or minor. If only ICD codes are available – which is often the case in large register studies – it is more difficult and sometimes nearly impossible to evaluate the severity of the malformation, notably when the code marks "other" or "unspecific" malformations, for example, ICD-10 codes Q55.8 = "Other specified malformations of male genital organs" or Q55.9 = "Congenital malformations of male genital organ, unspecified." A code indicating a ventricular septum defect (Q21.0) may represent a heart malformation which will need surgery or a defect which closes spontaneously and which will never play any role.

In the European congenital malformation-monitoring organization (EUROCAT), lists of malformations which should be regarded as major or minor are given. To use them, more detailed information is often needed than what is given by the standard ICD-10 code.

According to my view, the important thing is to eliminate minor malformations by removing from the analysis common and variably registered conditions in order to get a more stable concept. Since many years we have in Sweden used a method of excluding such conditions and we have called the remaining ones "relatively severe," in practice it will be rather similar to what is usually called major malformations, but among them will some minor conditions be left. It will, however, reduce the variability in recording as was evident above. The conditions which are excluded are preauricular tags, patent ductus arteriosus in preterm infants, tongue-tie, single umbilical artery, undescended testis, hip dysplasia, and nevus. These exclusions reduce the rate of malformations from about 5 to about 3 %. This leaves, for instance, preauricular pits which EUROCAT classifies as minor, but it is a condition which often needs surgery.

6.3 Single and Multiple Malformations

Infants may have more than one code for a congenital malformation. Experience has shown that teratogenic agents often give more than one malformation and sometimes give specific patterns of malformations. Infants with multiple malformations are therefore of a special interest. Two or more malformations in the same infant may have different causes.

Random hits are one explanation which may explain some infants with two malformations but very few infants with three or more malformations. If we suppose that 3 % of all infants have a major malformation, only one in 1100 will have two by the random occurrence of two different malformations and only about 1 in 37,000 will have three. The actual rates of infants with two or three different malformations are much higher.

Another explanation to the presence of two malformation codes for the same infant is the phenomenon of *sequences.* This means that one malformation is a

direct result of another, the primary one. Classical examples are that infants with spina bifida often have hydrocephaly or pes equinovarus as a result of the spinal cord malformation. Another example is pulmonary hypoplasia and facial dysmorphology as the result of absence of kidneys, Potter sequence. Infants with cardiovascular malformations often have more than one cardiovascular malformation code even though the primary damage to the heart rudiment probably is a single hit. Sequences should be regarded as single malformations, that is, as the primary anomalies which gave secondary changes.

A third group of multi-malformed infants are represented by *syndromes*. This term is often misused but should refer to defined constellations with a known cause. Examples are the rubella syndrome, with congenital cataract, hearing problems, and cardiovascular defects, and numerous genetic syndromes. If the cause is quite certain, such cases can be removed from the analysis but they are relatively few. At the detection of a new syndrome, often one leading malformation is first identified (in the case of the rubella syndrome, it was cataract and in the case of the thalidomide syndrome, it was amelia or phocomelia) and other characteristics are added later when groups of syndrome children are investigated.

The fourth group consists of constellations or patterns of malformations which are known but has no definite explanation, the *nonrandom association* of congenital malformations. Many such conditions are known; we can exemplify it with the VATER or VACTERL constellation which contains vertebrate, anal, trachea-esophageal, and radial or renal malformations (VATER), sometimes with cardiac malformations added (VACTERL). This is not a very unusual type of multi-malformed infant and may form an entity without known cause; it can be regarded as a syndrome looking for its cause! The explanation to a nonrandom association may be similarities in the embryogenesis and/or timing of the various malformations and may therefore not suggest a common cause like a drug exposure.

Finally we have the large group of multi-malformed infants which do not fit into any of the abovementioned groups. They may turn up to be unidentified syndromes or at least nonrandom associations when enough data have been collected. This group of infants is of great interest in a search for teratogenic drugs and should preferably be described in detail in the reports.

6.4 Causes of Congenital Malformations

6.4.1 Genetics

Some congenital malformations are monogenic conditions. Examples are achondroplasia which is usually a dominant mutation, some forms of microcephaly which are autosomal recessive, and some forms of hydrocephaly which are X-bound recessives. For these conditions, exposures during pregnancy are of little interest and such cases could be left out from analysis, but they are few. One also has to consider the possibility that the drug (if used before conception) could

cause a dominant mutation in the egg or sperm, resulting in a malformation. There is also a possibility that the drug causes a phenocopy, a condition which looks like a genetic condition. A classic example is warfarin which may cause a skeletal anomaly which resembles a genetic condition, chondrodystrophia calcificans. The constellation of malformations caused by mycophenolate mofetil (see p. 10) could be taken for the CHARGE association, a chromosome 18q deletion, or the HMC (hypertelorism-microtia-clefting) syndrome (Perez-Atyes et al. 2008). I think there is no reason to remove the very few cases of monogenic conditions which may turn up in an analysis. They may dilute the results but the effect will be small.

Many malformations have a genetic component which is more complex. So, for instance, orofacial clefts often occur in more than one family member, men with hypospadias have an increased risk to father a boy with this malformation, and couples who have had one fetus with spina bifida has a markedly increased risk to have another. Some investigators prefer to remove cases with a known family history of the malformation from the analysis, others do not. If the genetic trait is strong, such cases will dilute the material. On the other hand, it is possible that the genetic background makes the embryo especially sensitive for an environmental factor, for instance, maternal drug use. My preference is to keep cases with a family history; they might bias the risk estimate slightly toward null but this is not certain. If data are available, it is of course of interest to compare cases with and without a family history of the malformation, but it is rare that large enough numbers are present to allow such comparisons.

6.4.2 Chromosome Anomalies

A similar situation as with monogenic conditions exists for chromosome anomalies. The chromosome anomaly may be inherited from one of the parents or have occurred at the meiotic divisions at the formation of the egg or sperm. Many chromosome anomalies result in congenital malformations. If we take the most common autosomal anomaly, trisomy 21, it causes Down syndrome with typical dysmorphology and mental retardation but also with an increased risk for structural congenital malformations – heart defects – may occur in 40–50 % of these children and also other malformations occur in excess, e.g., duodenal atresia.

A difference between monogenic conditions and chromosome anomaly is that the latter are relatively common, even though modern prenatal diagnosis to some extent prevents the birth of such infants. The diagnosis is usually also definite after karyotyping. In analyses of specific malformations, infants with chromosome anomalies are usually excluded because the effect of the chromosome anomaly is such a strong cause of the malformation. There is, however, an interesting question – resembling the situation at familial malformations – that embryos carrying the chromosome anomaly could be more sensitive for external influences than normal embryos. Some studies have been made on Down syndrome to test this hypothesis but with no clear-cut results.

Chromosome anomalies should be included in the group "major malformations," but in the analysis of specific malformations, they should be left out (or treated as a separate group).

6.4.3 External Factors

Maternal use of drugs as a teratogenic factor is the main theme of the present book. Other external factors may disturb embryonic development with malformations as a result. If they somehow are associated with drug use, they may appear as confounders as will be discussed later. It is also possible that such factors may act synergistically with the drug.

Among such external factors can be mentioned some virus infections (notably rubella), strong ionizing irradiation, alcoholism, smoking, and some occupational exposures (Fixler and Threldkeld 1998). Some of them will be discussed in Chap. 8 on confounding. Other external factors are more uncertain like hot baths, showers or sauna, and nutritional deficits, and an association with specific drug use is also less likely.

6.5 Sources of Information on Malformations

In small studies information on malformations can be obtained by scrutiny of medical records from various disciplines. Sometimes interview or questionnaire information is obtained from parents or general practitioners who may be uncertain sources. For large-scale investigations one usually has to use register data.

There are different types of registers which can help to identify infants with malformations. There are specific malformation-monitoring registers around the world of varying quality and content. There are also international organizations which collect data from the various registers, e.g., the International Clearinghouse for Birth Defects Surveillance and Research and the European EUROCAT. Also within the USA, collaboration between different state registers occurs. An example is the National Birth Defects Prevention Study where data on some selected malformations are collected for epidemiological analysis from a number of state registers.

In the Scandinavian countries, medical birth registers exist which contain medical data on all pregnancies which end as deliveries. Late abortions are sometimes also included. Information on congenital malformations is incomplete when it is based on obstetric instead of pediatric information. In the Swedish register, infant information is given by qualified pediatricians who examine every infant born, but in spite of this, only a proportion of all malformations are identified. This information can be supplemented with discharge diagnoses from hospitalizations of the newborns and in some of the countries with data from specific registers of congenital malformation where reports are obtained from pediatricians and pediatric clinics. The linkage of the different sources of information is made with the use of the unique personal identification numbers of the mother and the infant. This system gives a relatively good ascertainment, but it is probably not complete. All internal

malformations are not detected in the newborn period, and follow-up is often only made during the first year of life when most but not all such conditions are identified.

The use of discharge diagnoses from neonatal units results in a complication. After some exposures, e.g., maternal use of antidepressants, neonatal morbidity increases and infants are often transferred to neonatal units, not because of a malformation but because of other morbidities. If no similar examination of non-transferred infants is made, a biased recording will be obtained. This may, for instance, explain the fact that most investigators find no effect on malformation rate after SSRI exposures (e.g., Källén et al. 2013), but in studies from Denmark (Pedersen et al 2009; Kornum et al. 2010; Jimenez-Solem et al. 2012), exclusively using discharge diagnoses from neonatal units, some malformation risks are seemingly increased.

Whichever technique for ascertainment is used, it is imperative that the same method is used for exposed and unexposed infants. If ascertainment is incomplete, it will reduce the power of the analysis but affect risk estimates only little as will be explained later in this book.

6.6 Prenatal Diagnosis and Induced Abortion

Today prenatal sonographic examination and other prenatal diagnostics are routine in developed countries. Then some malformed fetuses are identified and the woman can then choose to have her pregnancy interrupted with an induced abortion. The level of prenatal malformation detection depends on the equipment and the qualification of the investigator. In most but not all countries, there is an upper gestational age limit after which an abortion is not allowed. Detection of malformations at a second ultrasound around week 32 may then not result in an abortion. Late pregnancy detection, for instance of hydronephrosis, can increase the rate of registration after birth.

In some populations (e.g., Denmark, Finland) it is possible to link information on aborted fetuses with maternal drug use, in others (e.g., Sweden) law prohibits the registration of abortions with identification numbers and no linkage can be made.

Most malformations which are detected early enough to permit an abortion are relatively severe. A large proportion consists of chromosomally abnormal fetuses (which are of relatively little interest in studies of drug effects). Another large group is anencephaly and related malformations which are easily detected and nearly always aborted. This will result in the birth of only very few infants with anencephaly, and if analyses are restricted to infants born, an association between maternal drug use and anencephaly is difficult to detect. Other severe malformations are sometimes but not always detected and the fetus aborted. If aborted fetuses with spina bifida are not taken into consideration, an association with drug use can still be observed, but the study power will obviously be decreased. It can be illustrated with Swedish data on the association between maternal use of valproic acid and infant spina bifida. Among 5214 infants exposed in utero for anticonvulsants, 365

(7 %) were exposed to valproic acid. Four of the 5214 infants had spina bifida; all had been exposed to valproic acid. The expected number of spina bifida cases after valproic acid exposure is 0.28, and the observed number of four is significantly high (95 % Poisson confidence interval of 4 is 1.09–10.2). The risk estimate is a 14 times increase which agrees well with the 10–20 times risk increase stated in the literature. Prenatal diagnosis may have been intensified because of the valproic acid exposure which would have biased the risk estimate based on newborns toward null.

Another group of malformed infants which may be especially sensitive to prenatal diagnosis and may be a target for drug teratogenesis is multi-malformed infants. For this group a problem exists: aborted fetuses may be registered according to the malformation which was detected at the prenatal diagnosis, and other malformations present may not be recorded, notably if the aborted fetus was not autopsied by a fetal pathologist.

A study from Israel (Levy et al. 2012) claimed that exclusion of induced abortions biased the risk estimates toward null, illustrating it with data on folic acid antagonists. From the presented data, one can see that a majority of exposed neural tube defects in this population were detected and aborted (29 of 31), and this was true for about half of the cardiovascular defects (8 of 15). The corresponding percentages for unexposed cases were 15 % and 5 %, respectively. Thus the fact that women had used these drugs resulted in a considerable increase in induced abortions (as a result of intensified prenatal investigations) which made it nearly impossible to detect an effect on infants born, notably on neural tube defects. If this diagnostic increase was the same for well-known teratogens like anticonvulsants and notably valproic acid and for less known drugs is not clear from the study.

6.7 Grouping of Congenital Malformations

The concept of congenital malformations covers a large number of different conditions with different embryology. It is possible but unlikely that a teratogenic factor causes all types of malformations. There is a reason to divide the malformations into smaller and more homogeneous groups. This does not mean that the risk for any (major) malformation is uninteresting; this is actually the risk which the pregnant woman is mainly interested in.

There is no standard way to divide malformations into subgroups. The chapter division of the ICD code is often followed, but this is really not a good idea. The ICD codes were arranged in a way to make it easy to find a specific malformation and are therefore based on organ systems (with some exclusion like chromosomal anomalies or malformation syndromes). Each such group may consist of very different malformations with different embryogenesis. This will result in a grouping of malformations of different nature or to a spreading of related malformations to different groups.

Musculoskeletal malformations contain, for example, as different malformations as pes equinovarus and other positional foot defects, limb reduction defects, achondroplasia, and body wall defects. Even a subgroup of “body wall defects” contains

Table 6.1 Suggested grouping of congenital malformations among 5214 infants exposed to anticonvulsants in early pregnancy

Malformation	Number	Comment
Any malformation	365	
Relatively severe malformations	244	
Down syndrome	5	
Other chromosome anomalies	3	2 Turner syndrome
Neural tube defects	4	1 encephalocele, 3 spina bifida
Brain malformations	10	4 midline defects
Eye malformations	5	Different types
Orofacial clefts	19	12 cleft palate, 7 cleft lip/palate
Cardiovascular defects	92	51 only ventricular and/or atrium septum defect
Alimentary tract atresia	7	2 esophageal, 3 small gut, 2 anal atresia
Major kidney malformations	4	3 agenesis/hypoplasia, 1 cystic kidney
Hydronephrosis or urinary tract obstruction	12	9 hydronephrosis, 1 ureter obstruction, 2 vesico-ureter-renal reflux
Hypospadias	44	
Diaphragmatic hernia	4	
Body wall defects	2	1 omphalocele, 1 gastroschisis
Craniostenosis	5	
Poly- or syndactyly	18	11 polydactyly, 7 syndactyly
Limb reduction defects	3	1 cleft hand/foot, 1 absent leg, 1 longitudinal arm defect
"Syndromes"	2	1 probable Pierre-Robin sequence

very different malformations: omphalocele, gastroschisis, and large body wall defects, with different embryogenesis and epidemiological characteristics. Also limb reduction defects consist of different subgroups with different characteristics like transverse and longitudinal reduction defects.

Urogenital malformations contain very different types, for instance, absence of kidney or renal dysplasia, obstructive malformations leading to hydronephrosis, and hypospadias.

On the other hand, esophageal atresia, small gut atresia, and anal atresia show many similarities in embryogenesis and epidemiology but belong to different groups: Q39, Q41, and Q42.

Table 6.1 gives an example how an embryological more reasonable summary of a group of observed malformations can be given. It summarizes relatively severe malformations among 5214 Swedish infants exposed to anticonvulsants in early pregnancy, tentatively grouped according to embryological principles. Note that some infants had more than one malformation.

Ten of these infants had combinations of major malformations; five of them had hypospadias.

These cases should also be listed:

Malformations
Cleft palate + hypospadias + ASD + tongue malformation
Cleft lip/palate + hypospadias + VSD/ASD/CoA + syndactyly
Hypospadias + ASD
Hypospadias + VSD
Hypospadias + pes equinovarus
VSD/ASD + split hand and foot
VSD + polydactyly
Subaortic stenosis + bile duct atresia
Unspecified brain malformation + VSD
Malformation of anterior eye segment + ASD

ASD atrium septum defect, *CoA* coarctation of aorta, *VSD* ventricular septum defect

This type of reporting of malformations also makes it possible to add detailed materials from different investigations.

References

Bénthenod M, Frédérich A (1975) Les enfants des antiépileptiques. Pediatrie 30:227–248

EUROCAT. Malformation coding guides. http://www.eurocatnetwork.eu/aboutus/datacollection/guidelinesforregistration/malformationcodingguide

Fixler DE, Threldkeld N (1998) Prenatal exposures and congenital heart defects in Down syndrome infants. Teratology 58:6–12

Hansson JW, Myrianthopoulos NC, Harvey MAS, Smith D (1976) Risks to the offspring of women treated with hydantoin anticonvulsants, with emphasis on the fetal hydantoin syndrome. J Pediatr 89:662–666

Jimenez-Solem E, Andersen JT, Petersen M, Broedback K, Jensen JK, Alzal S et al (2012) Exposure to selective serotonin reuptake inhibitors and the risk of congenital malformations: a nationwide cohort study. BMJ Open 2:e001148

Källén B, Borg N, Reis M (2013) The use of central nervous system active drugs during pregnancy. Pharmaceuticals 6:1221–1285

Kornum JB, Nielsen RB, Pedersen L, Mortensen PB, Nørgaard M (2010) Use of selective serotonin-uptake inhibitors during early pregnancy and risk of congenital malformations: Update analysis. Clin Epidemiol 2:29–36

Levy A, Matok I, Gorodischer R, Sherf M, Wiznitzer A et al (2012) Bias towards the null hypothesis in pregnancy drug studies that do not include data on medical terminations of pregnancy: the folic acid antagonists. J Clin Pharmacol 52:78–83

Pedersen LH, Henriksen TB, Vestergaard M, Olsen J, Beck BH (2009) Selective serotonin reuptake inhibitors in pregnancy and congenital malformations. Population-based cohort study. BMJ 23:339: b3569. doi:10.1136/bmj.b3569

Perez-Atyes A, Ledo A, Bosso V, Sáenz P, Roma E, Poveda JL et al (2008) In utero exposure to mycophenolate mofetil. A characteristic phenotype. Am J Med Genet (Part A) 146A:1–7

Seip M (1976) Growth retardation, dysmorphic face and minor malformations following massive exposure to phenobarbitone in utero. Acta Paediatr Scand 65:617–621

Identification of Maternal Use of Drugs

7

Different methods have been used to identify women who have used drugs or specific drugs during pregnancy.

7.1 Questionnaires or Interviews in a Case-Control Setting

In many registers of congenital malformations, women who had a malformed infant are interviewed soon after birth about factors which could have caused the malformation like drug usage. In some, interviews are also made of normal controls, often next baby born with the same sex as the malformed infant. This is made in order to get a control material which is matched to the cases with respect to place, time of birth, and infant sex. As will be discussed later, infant sex can obviously affect malformation risk but in most cases not drug use during pregnancy. An advantage is that interviews are made a short time after delivery, but in spite of this a risk for recall bias exists.

In other studies, women are interviewed some time (sometimes a long time) after the delivery with similar interviews of controls. The largest such study is the US National Birth Defects Prevention Study (Yoon et al. 2001) from which numerous publications have appeared during the last decades. Exposure information is obtained from telephone interviews months after delivery. The nonresponse rate is high, about 30 %, which can also easily cause bias.

This type of data collection may be the only available in some settings like studies in developing countries, but results should always be looked upon with suspicion because of the risks of bias. In developed countries, other methods are preferable.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_7

7.2 Prospective Randomized Studies

The gold standard in medical research is the prospective randomized study. Such studies are rare in the field of drug teratogenesis for obvious ethical reasons. To this is added that such a study would with necessity be too small for an evaluation of rare outcomes like malformations. In one recent study, the authors performed a randomized study of the adverse drug reactions when doxylamine was used for nausea and vomiting in pregnancy (Koren et al. 2015). The study went on for 2 weeks, and 131 women got the drug and 125 got a placebo. The presence of congenital malformations in the infants was not studied, and it had been rather impossible to do so due to the low numbers. In this situation the condition which was treated was not very serious, and a placebo could be accepted.

A randomized study comparing two different but from the point of view equally effective drugs could be ethically acceptable. An example is a study on the suggested risk of paroxetine versus citalopram on cardiovascular defects. A paired analysis, supposing an 1 % cardiovascular disease rate in the total material, would need nearly 2300 pairs in order to demonstrate a doubling of the risk (alpha = 0.05, beta = 0.80). Even so, women may object to getting paroxetine as there is at least a suspicion that it may increase the risk for some cardiac defects.

From a practical point of view, a randomized study on teratogenicity of drugs is no reasonable option.

7.3 Teratology Information Services (TIS)

In many countries there exist TIS organizations. Doctors and sometimes patients can call the TIS and get information on possible risks associated with drug use during pregnancy. In most circumstances, the woman has already used the drug and wants information if the risk is so large that the pregnancy should be interrupted and perhaps also if she can continue using the drug during the rest of the pregnancy. This gives a perfect possibility to prospectively identify drug use, and at the conversation with the woman details like dose and timing of the use can be obtained and also information on other factors of interest, including other drug use than that which her concerns refers to.

There are two main problems with this approach. One is that it is difficult to get together large enough numbers in order to study malformation risks. The second problem concerns the follow-up procedure. This is usually made by contacts after the estimated time of delivery, either with the woman or with her doctor. The quality of this information can be debated, and one can notice that the reported malformation rates in controls are often quite low. A third problem is the obvious selection bias which is obtained as only data for women who are concerned about the drug use are included.

Controls are often women who have contacted the TIS with questions about exposures which are regarded as harmless, a procedure which is somewhat subjective.

One of the most active TIS organizations in the world is probably the Canadian Motherisk. A very large number of original papers have been published from this organization. Most are based on small materials. We can as an example take a relative recent study on gabapentin use which was based on 223 pregnancies exposed to gabapentin and 223 unexposed pregnancies, that is, the woman had asked for advice after an exposure which was judged non-teratogenic (Fuji et al. 2013). In the abstract it is said that the rates of major malformations were similar in the two groups ($p = 0.845$) and that the drug does not seem to increase the risk for a major malformation. This conclusion is based on 7 malformed infants among 170 live born infants after gabapentin and 5 among 201 live born infants in the control group. If these two rates are compared with Fisher exact test, the p-value is 0.40 (2 sided), and the odds ratio estimate is 1.68 (95 % confidence interval 0.45–6.85). Thus, the conclusion should be that the study is too small to exclude a nearly seven times increased risk, and the risk estimate shows a nearly 70 % increased risk. This is a typical situation when data are presented from one single TIS. The highest value of such a small study is that it can exclude a really strong teratogenic effect of the magnitude which was seen after thalidomide where about 20 % of exposed infants had severe malformations and these were of a similar type.

By adding data from different TIS, it has been possible to increase numbers (e.g., McElhatton et al. 1996). Such a combined material runs a risk of heterogeneity both in the identification of the drug use and in the method and completeness of the follow-up of the pregnancies, and this complication has to be taken into consideration when data are analyzed. It is also important that women acting as controls are selected proportionally between the involved systems so exposed cases from different countries are not compared with a control material from one country.

Sometimes no control material is presented, but the recorded rate of malformed infants is compared with a general figure of "1–3 %." It should be remembered that controls are needed not so much in order to estimate the true rate of malformed infants in the population studied but more importantly to evaluate how complete ascertainment has been. Ascertainment rates probably vary more than actual rates.

7.4 Prospective Studies of Pregnant Populations

In an ideal world, there would exist recorded exposure data for all pregnancies in a population. One way to obtain that is to start a research project when all pregnant women are questioned early in pregnancy (in order to avoid recall bias) about drug use since she became pregnant. Questions on other variables, e.g., pregnancy history, smoking, and BMI, could also be included. Then all pregnancies should be followed until delivery, and the infants born should be examined for morbidity, including congenital malformations. Such studies have usually been made as time-limited research project. One of the earliest such study was the Collaborative Perinatal Project which was first mounted in the 1950s (thus before thalidomide) and collected data from 14 university-connected hospitals in USA (Heinonen et al. 1977). This study covered over 50,000 pregnancies and gave important

information on drugs used at that time (1958–1965). The authors stressed in their analysis that some malformations have roughly the same rate in the participating hospitals, while other malformations vary strongly in rate. They also distinguished between major and minor malformations and made numerous detailed tabulations of rates of groups of malformations, mainly divided after organ systems. This study is still an important source of information, but naturally modern drugs were not included.

After the thalidomide tragedy, a number of smaller prospective studies were carried out. One of these was made at the Malmö hospital in Sweden in 1963–1964. All pregnant women who attended prenatal care and the only hospital in the city were interviewed repeatedly during pregnancy, among other things concerning drug use and smoking. It was a small study of only 6,300 pregnancies, and numbers were too small to investigate congenital malformations in detail. One report from the study concerned maternal smoking (Kullander and Källén 1971), and further reports described some common drug groups (Kullander and Källén 1976).

More recently prospective studies have been performed in Denmark and in Norway. The Norwegian study (MoBa) is based on questionnaires sent to the mothers in connection with their attendance to ultrasound investigations around week 15 (Magnus et al. 2006). The study was made during 1999–2005. The participation rate is only about 43 %, but the study contains detailed information on more than 64,000 pregnancies. The study has been used for many different studies including some on drug use during pregnancy. The Danish National Birth Cohort study was performed in 1996–2002 and referred to more than 100,000 births, approximately 30 % of the women who gave birth (Olsen 2001; Liew et al. 2014). Information on drug use in pregnancy was obtained by telephone interviews at gestational weeks 12 and 30.

7.5 Pregnancy Registers

This term is often used for registers of pregnancies exposed for specific drugs and have usually been organized by the drug industry. An example is the GlaxoSmithKline Lamotrigine Pregnancy Registry. This register "is intended to provide an early signal of potential risks in advance from formal epidemiological studies" (Foreword to Interim Report from Lamotrigine Pregnancy Registry, 2004). Prospective data on drug exposure were collected from a number of countries around the world. No control material was available, but the observed rates of congenital malformations were compared with data in the literature.

The register also collected retrospective data, after the outcome of the pregnancy was known. Such data are probably biased but can give information on an aggregation of a specific type of malformations.

The register closed down after 18 years (Cunnington et al. 2011) and had then collected 1558 first trimester monotherapies, reported prospectively without finding any signs of an increased risk for any specific congenital malformation (based on 35 malformed infants).

Another example on a register which is specifically built for the study of one group of drugs is the Massachusetts General Hospital National Pregnancy Register for Atypical Antipsychotics (Cohen et al. 2015). It consists of a prospective recording of women using such drugs in pregnancy and controls, mainly women with psychiatric conditions but using other drugs. The register is still small: 353 women using such antipsychotics and 134 controls and the power to study congenital malformations is low, to which is added that the ascertainment of malformations appears low, less than 2 %.

7.6 The Swedish Medical Birth Register

In the Nordic countries, there are medical birth registers which contain data on (nearly) all deliveries in the countries. The Swedish register was started in 1973, and information on drug use was added in 1994. Since 1982, the register is based on copies of the medical documents at the prenatal care (which practically all pregnant women attend and which is free of charge), delivery, and the pediatric examination of the newborn. Identical medical forms are used in all Swedish hospitals since 1982. This form contains space for the recording of drugs used as reported by the woman at the first prenatal care visit, usually in week 10–12. Later drug use as initiated by the prenatal care is also recorded. This system has formed a large data base with drug exposure information obtained in early pregnancy, at present containing data for 1.7 million deliveries. It has been used in numerous studies on the effect of maternal drug use on infant outcome. The drug information is stored as Anatomic, Therapeutic, Chemical (ATC) codes. There are some information on dosage and timing, but these are incomplete and cannot be used in most instances (Källén and Otterblad 2001).

In practice this is equivalent with an ongoing prospective study even though it is less detailed than most prospective studies of drug use during pregnancy.

The information on drug use is thus based on what the woman tells the interviewing midwife. She may have forgotten about temporarily used drugs or may avoid telling about the use of some "sensitive" drugs, for instance, use of recreational drugs. Furthermore, the midwife may miss to record the reported drug or may do it in a way which makes it difficult or impossible to interpret the recording. Drug names are recorded in clear text and then centrally transferred into ATC codes, a partly rather tedious work, notably when drug names are wrongly spelt or written down in a hard way to read. For these reasons, some drug exposures will be missed. It is also possible that the woman mentioned drugs which were used outside the period of pregnancy – an example is that some women report drugs used for ovulation stimulation in spite of the fact that they were used before pregnancy. If the woman's first visit to prenatal care is early, e.g., in weeks 6–7, she may have used drugs later during the first trimester which were not recorded. An effort to compare prescription data with data from the Swedish Medical Birth Register indicated a good agreement for chronically used drugs, less good for drugs used temporarily

(Stephansson et al. 2011). We will come back to the significance of missing data and inclusion of invalid data.

The system thus has some disadvantages but also advantages: the data base is growing, and newly introduced drugs will find their way into it while most prospective research studies are time limited. Information on drug use is retrieved in early pregnancy, and a recall bias is unlikely even though one cannot exclude the possibility that, for instance, a bleeding in early pregnancy may affect the information given.

7.7 Prescription Registers

During the last decades, prescription registers have appeared in some countries including the Nordic ones. These registers are formed from the computerization of prescriptions at the drug stores when a patient fills a prescription. Thanks to the use of personal identification numbers such information can be linked to registers over pregnancy outcome. This gives a relatively simple access to information on which drugs the woman has been prescribed just before or during pregnancy. As pointed out above, the registers can also be used to identify diseases in the child in long-term follow-up studies.

There are some drawbacks of such systems. One is that drugs given in association with hospitalizations will not be included; neither is over the counter drugs. A more serious problem is that it is not known if the woman actually used the drug after buying it. Especially if she gets a prescription when she is early pregnant – or even only suspects she is pregnant – she may be unwilling to take the drug or postpones treatment until later when she has passed the most dangerous period, the first trimester. This will include unexposed pregnancies in the group thought to be exposed. On the other hand, she may well have used drugs which she got on prescription even years before the pregnancy – most drugs have a long shelf-life. She may also have used drugs which she had obtained in other ways, from partner, friends, or bought via the Internet. Data from prescription registers will therefore have uncertainties both by including or excluding actual exposures.

Relatively few studies have been made on the validity of such data. The problem with drugs used during hospital care was illustrated in studies from Sweden (Linder et al. 2015) and from Denmark (Haerskjold et al. 2015), both studying palivizumab use as prophylaxis for RSV in high-risk children.

One study in Norway used the prescription register as the gold standard and studied drug information in the Medical Birth Register (Espnes et al. 2011). For most drug categories, the latter data were much less complete than the prescription data. So, for example, there were a total of 701 instances recorded with diazepam, 612 of them only in the prescription data. This probably shows a low registration in the birth register but a noncompliance can also explain the discrepancy. Among the 89 such cases which were reported in the birth register, 26 (29 %) were not identified from the prescription register. For anticonvulsants, the birth register identified

497 exposures, 51 of which (10 %) were not found in the prescription register – and 290 were only identified in the latter register.

A comparison has been published of data on antidepressant use from a prescription register and interview data from the Swedish Medical Birth Register (Källén et al. 2011). The results indicated that for studies of exposures during early pregnancy, interview data gave a more complete and correct picture than prescription data – if the latter should be used, they should be limited to prescriptions given during pregnancy or possibly include prescriptions given within 1 month before pregnancy. If we accept that exposure actually occurred if the woman had reported so or if she had got a prescription in months 2–3 of pregnancy ($n = 5750$), 78 % were identified from the interview studies and only 55 % from prescriptions.

During the second to third trimester, no interview data were available, only information from the prenatal care medical records that the woman had been instructed to take an antidepressant – she may or may not have got a prescription from the prenatal care or from another source or could already have access to the drugs. In this situation, data from prescriptions were more complete as judged from the effect on neonatal complications. This indicates that women who had been recommended to take antidepressants to a large extent did not follow the advice or she used medicines she already possessed. Against the latter possibility speaks the fact the women who got no prescription but had been recommended by the prenatal care to take the medicine had no increased risk of complications (OR = 0.85).

This exercise demonstrates some problems to use prescription registers in order to identify drug use among pregnant women.

7.8 The Effect of Errors in Drug Exposure Ascertainment

As mentioned above in this chapter, some methods of drug ascertainment may error by including non-exposed cases as exposed. Non-exposure may be the result of the use of prescription registers as source of ascertainment or be due to misinformation in interviews. It may also be due to the use of drugs outside the sensitive period for the malformation studied. An effect of this error will be a bias of risk estimates toward null, and it cannot result in too high-risk estimates. It may give a false impression of harmlessness of a drug if the risk estimate is too low and may therefore not reach statistical significance.

If a woman had used a drug and this was not registered, it hardly affects risk estimates but will of course reduce the power of the study. It will mean that among the women regarded as unexposed some were actually exposed, but they will be few compared with the truly unexposed women and therefore hardly affect the risk estimate for the unexposed women.

Let us take a hypothetical situation where in a population of 100,000 women, 1000 had used a drug with a three times increased risk for a congenital malformation and the background rate of malformations in the unexposed women was 3 %.

If all 1000 exposed women were identified, we would have found 90 infants with malformations, a risk ratio of 3.0, 95 % CI 2.4–3.6.

If instead we had a group of 1000 women, only half of which were actually exposed, the number of malformed infants would be 60 (45 + 15, 6 %)and among the remaining 99,000 unexposed women, of course the risk would still be 3 % and the risk ratio would be 2.0, 95 % CI 1.5–2.5.

If only half of the 1000 exposed women (500) were identified, we would have 45 malformed infants, and among the apparently unexposed 99,500 women, there would be 3015 malformed infants (2970 + 45), a risk ratio of 2.97, 95 % CI 2.2–3.9: a very slight decrease of the risk ratio estimate but a wider confidence interval.

The most common way to get a falsely high-risk estimate is to rely on retrospectively ascertained exposure information. Irrespectively of how trained the interviewer is or how well formulated the questionnaire is, it is difficult to avoid the fact that a woman who has had a malformed child is more likely to remember – or even falsely recall – the use of drugs than a woman who has a normal child. This is one of the two main drawbacks of retrospective case-control studies using healthy controls. The second is the sometimes high rate of nonresponders (often about 30 %) which can give strongly biased results. It is sometimes argued that the fact that statistically significant results are sometimes obtained for some drugs but not for others argues for a true effect. A likely interpretation, however, is that for all studied malformations, the risk estimates scatter around a common increased estimate, which likely is due to recall and/or nonparticipation bias.

7.9 Information on Dosage of Drugs Used

Obviously the amount of the drug used is of interest but perhaps not so much as is often stated. There are situations when a dose dependency has been described, e.g., for anticonvulsants (Tomson et al. 2011) and paroxetine (Bérard et al. 2007). Quite often the information on dosage is obtained from prescription information, and it is far from certain that the prescribed dose is identical with that taken. Furthermore, the value of interest is hardly the dose taken by the woman but the amount which reaches the embryo. Different metabolisms in different women may result in different concentrations reaching the embryo after the same amount of drug taken. Up to 25 % of commonly prescribed drugs are metabolized by a highly polymorphic hepatic enzyme CYP2D6. Given the same dosage, phenotypic slow metabolizers of CYP2D6, that is, 7–10 % of a Caucasian population (Bertilsson et al. 2002), may show a higher serum concentration than extensive metabolizers, and this may increase the risk for malformations.

When only single tablets have been taken, the risk for damage is obviously lower than when a drug has been used for a longer period. This will mean that inclusion of cases who took only a few tablets in the overall risk estimate may bias it toward null in a similar way as inclusion of cases with exposures outside the sensitive embryonic period.

References

Bérard A, Ramos É, Rey E, Blais L, St-André M, Oraichi D (2007) First trimester exposure to paroxetine and risk of cardiac malformations in infants. The importance of dosage. Birth Def Res (Part B) 80:18–27

Bertilsson L, Dahl ML, Dalén P, Al-Shurbaji A (2002) Molecular genetics of CYP2D6: clinical relevance with focus on psychotropic drugs. Br J Clin Pharmacol 53:111–122

Cohen LS, Viguera AC, McInerney KA, Freeman MP, Sosinsky AZ, Moustafa D et al (2015) Reproductive safety of second-generation antipsychotics: current data from the Massachusetts General Hospital National Pregnancy Register for atypical antipsychotics. Am J Psychiat. doi:10.1176/appl.ajp.2015.150400506

Cunnington MC, Well JG, Messerheimer JA, Ferber S, Yerby M, Tennis P (2011) Final results from 18 years of the International Lamotrigine Pregnancy Registry. Neurology 76:1817–1823

Espnes MG, Bjørge T, Engeland A (2011) Comparison of recorded medication use in the Medical Birth Registry of Norway with prescribed medicines registered in the Norwegian Prescription Database. Parmacoepidemiol Drug Saf 20:243–248

Fuji H, Goel A, Bernard N, Pistelli A, Yates LM, Stephens S et al (2013) Pregnancy outcomes following gabapentin use. Neurology 80:1565–1570

Haerskjold A, Henriksen L, Way S, Malham M, Hallas J, Pedersen L et al (2015) The Danish National Prescription Registry in studies of a biological pharmaceutical: palivizumab – validation against two external data sources. Clin Epidemiol 7:305–312

Heinonen OP, Slone D, Shapiro S (1977) Birth defects and drugs in pregnancy. Publishing Sciences Group, Inc., Littleton

Källén B, Nilsson E, Otterblad OP (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

Källén B, Otterblad OP (2001) Monitoring of maternal drug use and infant congenital malformations. Does loratadine cause hypospadias? Int J Risk Safety Med 14:115–119

Koren G, Clark S, Hankins GD, Caritis SN, Umans JG, Miodovnik M et al (2015) Maternal safety of the delayed-release doxylamine and pyridoxine combination for nausea and vomiting in pregnancy; a randomized placebo controlled trial. Pregnancy Childbirth 15:59. doi:10.1186/s12884-015-0488-1

Kullander S, Källén B (1971) A prospective study of smoking and pregnancy. Acta Obst Gynecol Scand 50:83–94

Kullander S, Källén B (1976) A prospective study of drugs and pregnancy. 1. Psychopharmaca. Acta Obst Gynecol Scand 55: 25–33; 2. Anti-emetic drugs. Acta Obst Gynecol Scand 55: 105–11; 3. Hormones. Acta Obst Gynecol Scand 55: 221–4; 4. Miscellaneous drugs. Acta Obst Gynecol Scand 55:287–295

Liew Z, Ritz B, Rebordosa C, Lee P-C, Olsen J (2014) Acetaminophen use during pregnancy, behavioral problems, and hyperkinetic disorders. JAMA Pediatr 168:313–320

Linder M, Byström C, Keiler H, Bergman G, Haerskjold A (2015) Use of palivizumab is underestimated in the Swedish Prescribed Drug Register – implications for register-based drug studies. Clin Epidemiol 7:45–51

Magnus P, Irgens LM, Haug K, Nystad W, Skjærven R, Stoltenberg C, and the MoBa Study Group (2006) Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol 35:1146–1150

McElhatton PR, Garbis HM, Eléfant E, Vial T, Bellemin B, Mastroiacovo P et al (1996) The outcome of pregnancy in 689 women exposed to therapeutic doses of antidepressants. A collaborative study of the European Network of Teratology Information Services (ENTIS). Reprod Toxicol 10:285–294

Olsen J (2001) The Danish National Birth Cohort – its background, structure and aim. Scand J Publ Health 29:300–307

Stephansson O, Granath F, Svensson T, Haglund B, Ekbom A, Kieler H (2011) Drug use during pregnancy in Sweden – assessed by the Prescribed Drug Register and the Medical Birth Register. Clin Epidemiol 3:43–50

Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Sabers A et al for the EURAP study group (2011) Dose-dependent risk of malformations with antiepileptic drugs: an analysis of data from the EURAP epilepsy and pregnancy register. Lancet Neurol 10:609–617

Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SJ et al (2001) The national birth defects prevention study. Public Health Rep 116(Suppl):32–40

The Problem of Confounding

8

Confounding is obtained if a factor affects both exposure risk and outcome risk.

A confounder will affect the strength of the effect of the exposure on the outcome or may even completely explain it. The effect of the confounder should therefore be removed as well as possible in the analysis. We will come back to how this can be done.

How do we know which confounders are involved? We often do not know all possible confounders, or we cannot measure them. It is therefore always a possibility that a noticed effect is caused by unidentified confounders or confounders which we have not been able to adequately take into consideration. In these discussions one has to apply some common sense. Is it reasonable that the factor in question affects the rate of the outcome and also the exposure rate in a way which will affect the risk estimates? Both effects are necessary for the identification of a confounder.

8.1 Mediating Factors Should Not Be Adjusted for

Quite often one sees that adjustments have been made for factors which are no confounders. Often this does not matter much, but a definite exception is adjustment for mediating factors.

A mediating factor is affected by the exposure and is then causing an outcome. So, for example, some drugs increase the risk for preterm birth, and preterm birth results in an increased risk of neonatal morbidity.

Exposure → Mediating factor → Outcome

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_8

If this is the only way the exposure affects the outcome, adjustment for the mediating factor will eliminate the effect completely. In one study, an increased risk for childhood and young adult cancer was seen in children conceived by in vitro fertilization, IVF (Källén et al. 2010), and it was suggested that this was a consequence of the well-known increased risk for preterm birth, neonatal asphyxia, and low Apgar score after IVF, which would therefore act as mediating factors. In another study on partly the same material, one adjusted for preterm birth (and congenital malformations) and no increased risk of cancer was then found (Jerhamre Sundh et al. 2014). The conclusion to be drawn is not that there is no increased cancer risk among infants born after in vitro fertilization, but that it is mediated via an increased risk for preterm birth or factors associated with preterm birth, as previously suggested.

Adjustment for a mediating factor is useful if one wants to see if the effect of an exposure is completely or partly due to the mediating factor but should only be made for that purpose. It is then useful to make separate analyses with and without adjusting for the mediating factor. If one stratifies for preterm and term births and no effect remains in either stratum, one has shown that the effect of the exposure is completely due to preterm birth – if an effect remains within one or both strata, it seems likely that the effect does not (only) act via the studied mediator. In this example it should be remembered that preterm birth can mean anything from extreme preterm birth (e.g., <27 weeks) to late preterm birth (e.g., 35–36 weeks), so a remaining effect among preterm births could be due to different distributions of gestational length. In that situation one would have to adjust for actual pregnancy length.

8.2 Some Common and Sometimes Important Confounders

8.2.1 Year of Birth

When data are collected during a long period, one should consider if the registered rate of the studied outcome varies with year of delivery and if the use of the drug under investigation varies with year of delivery. This is exemplified in Fig. 8.1 which shows data for antidepressant use and the occurrence of septal heart defects in Sweden.

One can see that the registration (and perhaps use) of antidepressants has been increasing up to 2012, followed by a drop during 2013, perhaps associated with an change in the administration of the register. The rate of diagnosed or registered cases with ventricular or atrium septum defects (without other heart defects) has declined during these years. If adjustment for year of delivery is not made, the association between the exposure and the outcome could be affected; in this case the result seemed to be a reduction of the estimate because the rate of the exposure increases and the rate of the outcome decreases. This graph is based on 9 873 infants exposed to antidepressants in early pregnancy among 971 287 infants studied – only 67 had a septum defect after antidepressant exposure. The crude odds ratio for a

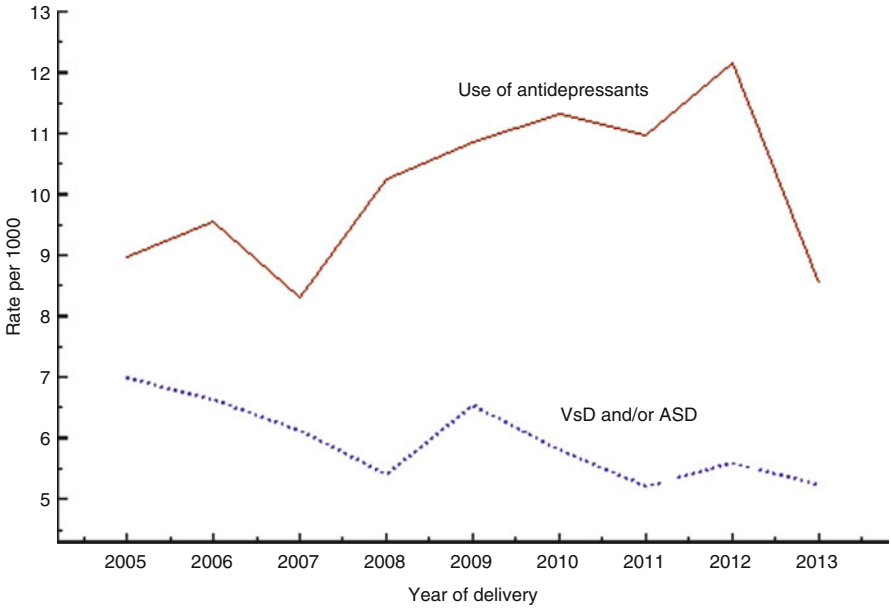


Fig. 8.1 Diagram showing rates of antidepressant use and of cardiac septum defects. Sweden, 2005–2013

septum defect after antidepressant exposure was 1.09, and it increased to 1.15 if adjustment was made for year of birth, thus nearly a doubling.

The adjustment for year of delivery is obviously very important when long-term effects are studied because of the differences in follow-up time.

8.2.2 Maternal Age

Drug use during pregnancy varies much according to maternal age. Some examples are given in Fig. 8.2. For many drugs use increases with maternal age, but for some it decreases, e.g., antibiotics.

For some congenital anomalies, notably chromosome anomalies, a strong relationship with a high maternal age is seen. For a few malformations, a high rate is associated with young maternal age (e.g., gastroschisis), while for the majority the age dependency is relatively weak (e.g., cardiovascular defects, Fig. 8.3). Note the markedly different maternal age dependencies of gastroschisis and omphalocele, both abdominal wall defects.

For other outcomes than congenital malformations, maternal age plays a more important role, for instance, for preterm births (see Fig. 5.1).

When adjustments are made for maternal age, 5-year age groups are often used. When the risk is steeply changing with maternal age, for instance at Down syndrome or gastroschisis, adjustment for 1 year is recommended.

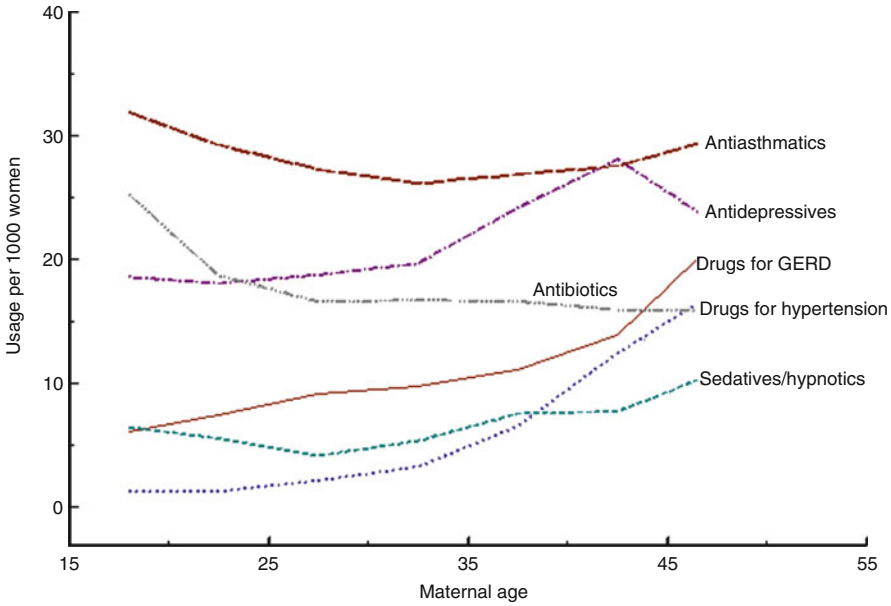


Fig. 8.2 Use of some drug groups at different maternal ages

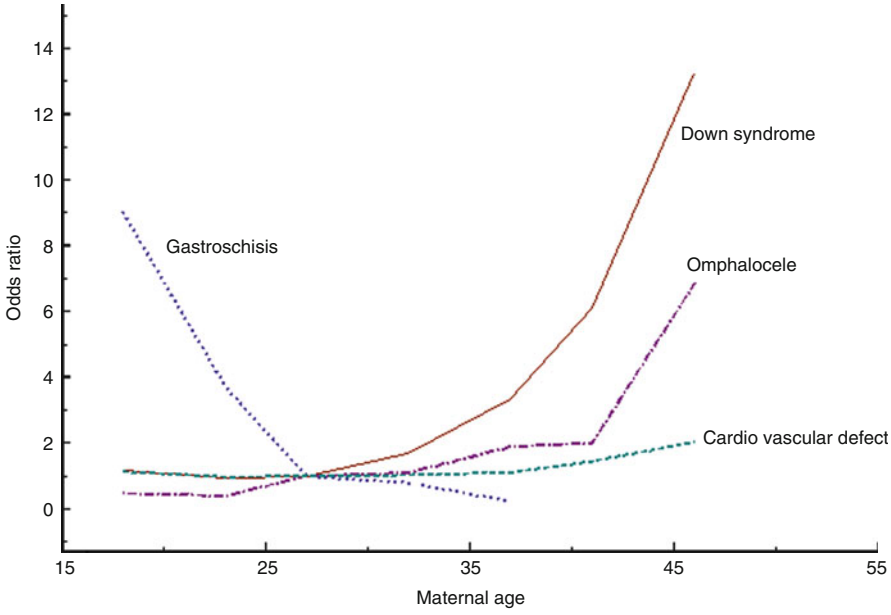


Fig. 8.3 Odds ratio for some malformations according to maternal age, adjusted for year of birth

8.2.3 Parity and Gravidity

These two concepts – number of births and number of pregnancies, respectively – are of cause strongly connected. The difference is made up of spontaneous or induced abortions and of extrauterine pregnancies. Among these outcomes, perhaps previous spontaneous abortions are the most likely one as confounders for outcome of a later pregnancy (see below). Instead of working with the concept gravidity, I prefer to use parity and number of previous spontaneous abortions as separate variables.

Parity is the number of previously born infants. A pregnant woman expecting her first child is a nullipara (parity 0) during pregnancy, but after delivery she is of parity 1. This is a bit confusing, and it is useful to define what one means with parity. In developed countries rather few women are of higher parities than four. Therefore, in analyses four and higher parities are often grouped together.

Drug use is often higher at parity 1 than at higher parity, if adjustment for maternal age is made. The occurrence of most malformations is not influenced by parity, but for some, e.g., cardiovascular defects and hypospadias, the risk is higher at first parity than at higher parities after adjustment for age (Källén 2014).

8.2.4 Smoking

Information on smoking is usually obtained by interview or questionnaire. Such information may be somewhat uncertain, but repeated tests of the data in the Swedish Medical Birth Register indicate a reasonably good quality. Generally, one finds similar effects of smoking registered prospectively or retrospectively. Smoking is often stated as no/yes, but it is useful if some quantification can be made, notably in populations where smoking is relatively prevalent. In the Swedish Medical Birth Register and in some other sources of information, one has chosen to distinguish between smoking less than 10 cigarettes per day and 10 cigarettes or more per day. The latter group is often not very large but will also contain smokers of 20 cigarettes per day or more, and this subgroup may vary between different exposure groups.

Smoking is associated with a high use of some drugs (Fig. 8.4), notably psychotropic drugs,

The rate of some malformations is influenced by maternal smoking (Källén 2002; Hackshaw et al. 2011), and there is a statistically significant effect of maternal smoking on the occurrence of any major malformation (Källén 2014). Figure 8.4 shows the adjusted odds ratio for smoking at some groups of congenital malformations.

There is an effect of maternal smoking on preterm birth and notably on intrauterine growth retardation. Also other ill effects of maternal smoking have been described (Källén 2001).

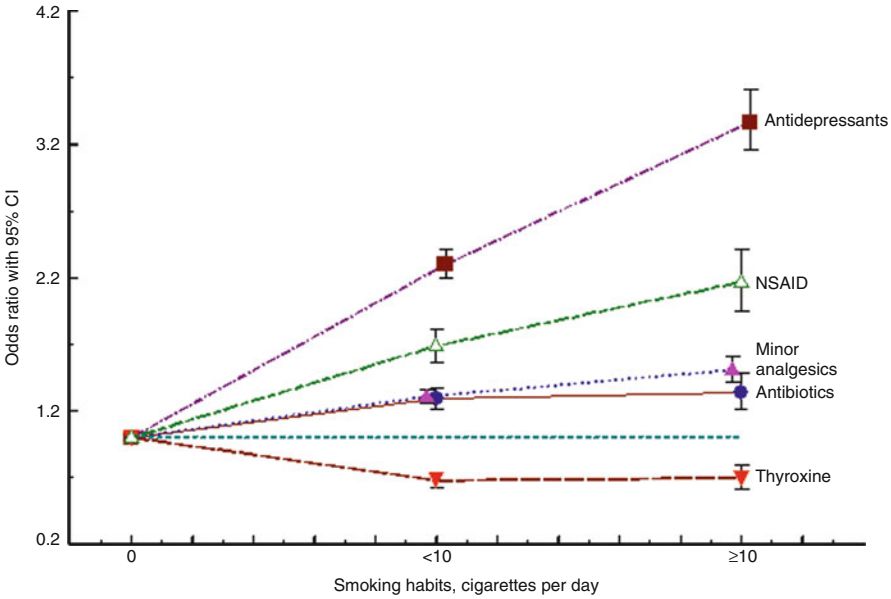


Fig. 8.4 Smoking habits among women using different drug groups in early pregnancy. ORs are adjusted for year of birth, maternal age, parity, and BMI

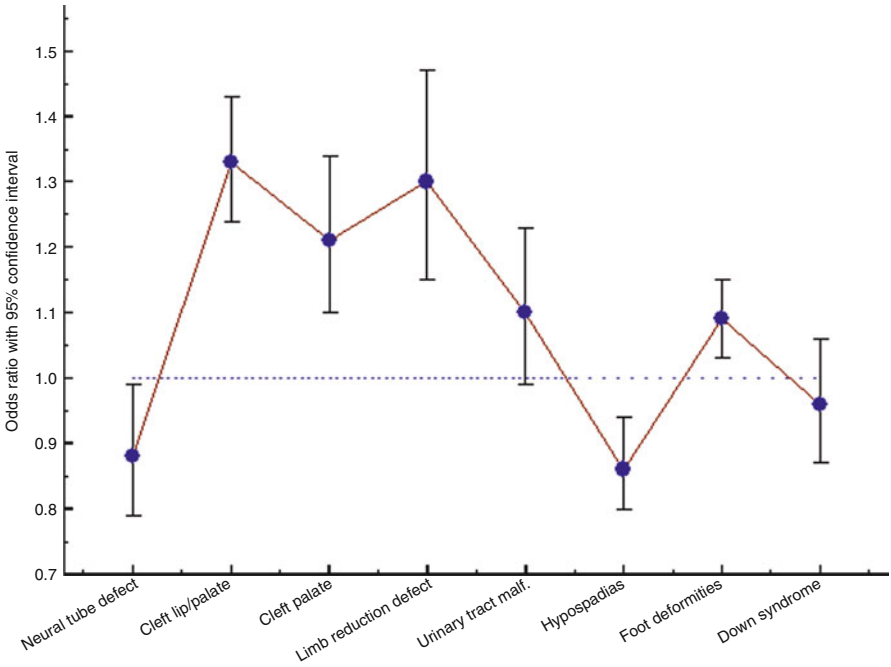


Fig. 8.5 Adjusted odds ratio for maternal smoking for some groups of congenital malformations (After Källén (2002))

8.2.5 Use of Alcohol

Maternal alcoholism or high use of alcohol is certainly a risk factor for many pregnancy outcomes, including congenital malformations. The result of maternal alcoholism is a typical condition, the fetal alcohol syndrome (FAS), characterized by intrauterine growth retardation, certain dysmorphic features, an effect on cognitive development, and an increased risk for some specific malformations, notably cardiovascular defects (Jones et al. 1973). The use of moderate amounts of alcohol during pregnancy should be avoided but no clear-cut effects are seen, not even on preterm births (Smith et al. 2015). The use of alcohol and the presence of alcoholism differ between different societies.

The use of alcohol during pregnancy is sometimes difficult to ascertain. Interview data are often unreliable because women may be concerned about stating their alcohol habits. In Sweden women coming to the first prenatal visit are interviewed by a midwife with a standard form which also includes questions on alcohol use. Very few women have then spontaneously admitted the use of alcohol during pregnancy. It is important to identify women with severe alcohol problems, but a simple question is hardly useful. Other methods are needed for their identification.

8.2.6 Overweight and Obesity

Overweight and obesity are identified and registered as prepregnancy BMI. If the woman comes early in pregnancy to prenatal care, BMI may be based on present body weight and height; otherwise, one has to rely on reported prepregnancy weight which may be uncertain.

The usual division of BMI is into <18.5 (underweight), 18.5–24.9 (normal), 25–29.9 (overweight), 30–34.9 (moderate obesity), 35–39.9 (severe obesity), and ≥40 (morbid obesity).

Overweight and obesity are associated with the use of some drug categories. Figure 8.6 shows the effect of maternal BMI on the use of three drug groups. The effect of BMI on the use of antibiotics is rather weak (although statistically significant), while the effect on the use of minor analgesics and antidepressants is marked. For most drug groups, usage increases with BMI.

Maternal obesity has been shown to have important effects on various pregnancy outcomes, including many congenital malformations (Blomberg and Källén 2010). Figure 8.7 summarizes the effect of maternal obesity on the risk for a number of congenital malformations.

Also other pregnancy complications are associated with maternal obesity (Cedergren 2006; Mission et al. 2015).

8.2.7 Subfertility

Subfertility means difficulty to get pregnant. The time it takes to achieve a pregnancy with unprotected intercourse is measured as the “time to pregnancy,” TTP.

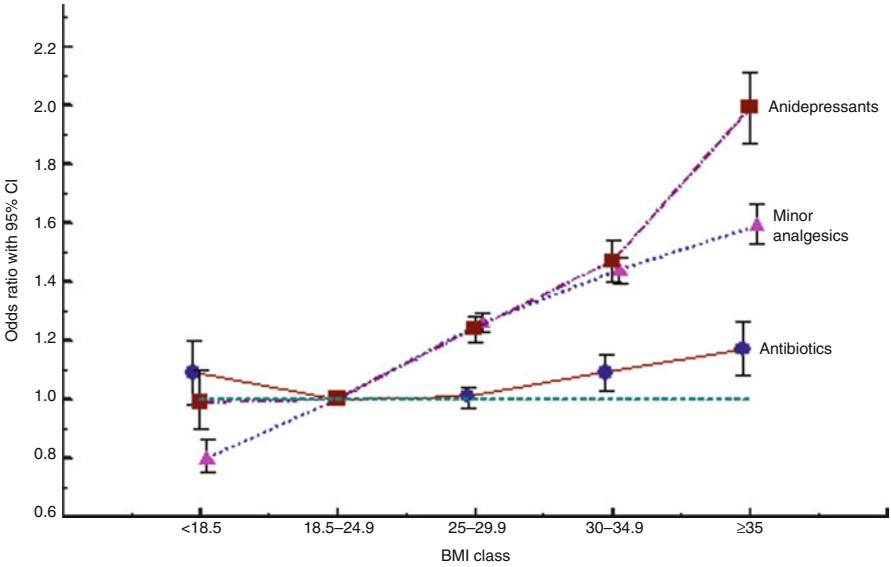


Fig. 8.6 Maternal use in early pregnancy of three drug groups according to maternal BMI. Adjustment for year of birth, maternal age, parity, and smoking

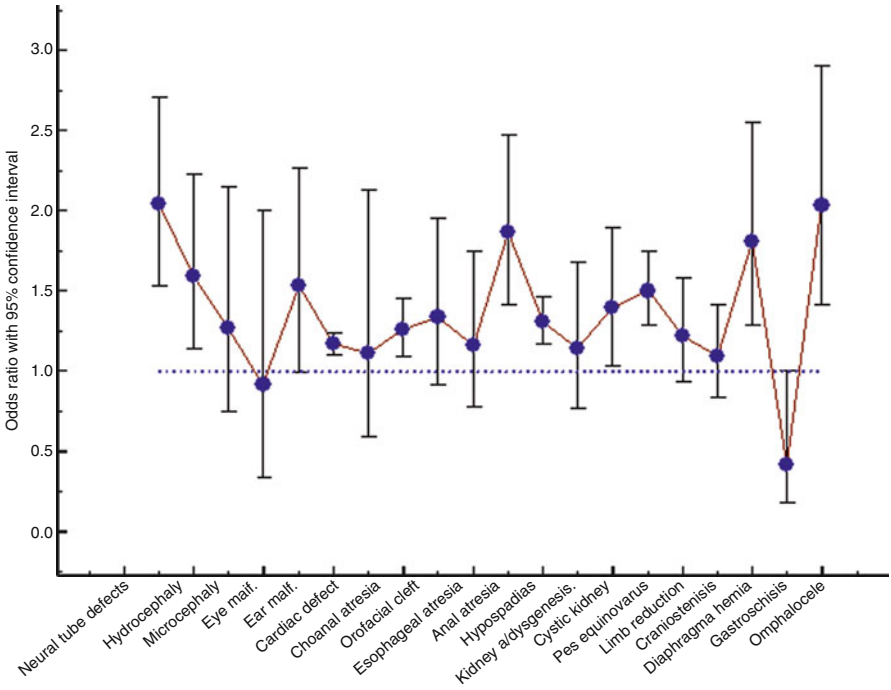


Fig. 8.7 Adjusted odds ratio for maternal obesity at different infant congenital malformations (Blomberg and Källén 2010)

A common definition of subfertility is that the couple should have tried to achieve a pregnancy at least 1 year without success. Even then, a considerable proportion of the couples will have a spontaneous pregnancy after further time, but perhaps 5 % will not (Gnoth et al. 2005). They can then be regarded as infertile, and their only chance to get a pregnancy is by medical treatment, including in vitro fertilization (IVF).

Information on TTP relies on information from the couple and will be sensitive for various biases including recall bias. For epidemiologic purpose, one will have to do with that information, imperfect as it may be. Many studies have shown that subfertility is associated with an increased risk for pregnancy and delivery complications, and this seems to be the major cause of abnormalities (including congenital malformations) seen after IVF. Subfertility is obviously associated with the use of fertility drugs but may have an effect also on the use of other groups of drugs, either an increased usage (e.g., thyroxin) or a reduced usage (e.g., sedatives/hypnotics and antidepressants) (Källén 2009).

8.2.8 Race/Nationality

In many populations, racial differences in pregnancy outcome are important, partly because of associations with socioeconomic conditions, partly because of genetics. This is true also for some congenital malformations, e.g., neural tube defects which are much more frequent in the US population among whites than among blacks. In other countries, e.g., the Scandinavian, the race issue is much less important – and furthermore in some countries it is illegal to register race in official registers. A proxy for race in the Scandinavian countries is the country of birth of the parents. In an analysis of the years 1978–1993 (Källén 1998), it was shown that the major anomaly in delivery outcome in Sweden according to maternal country of birth was a possible worse outcome for women who came from sub-Saharan Africa. At that time the proportion of these women was relatively low, less than 0.1 % in 1978 and about 1 % 1993. For the period 2005–2012 (Fig. 8.8), the proportion of women who were born in sub-Saharan Africa increased from 1.5 to 3 %. The majority of these women were probably black. The proportion of women born in East Asia increased from 0.2 % in 1978 to 1.4 % in 1993 and from 2.1 to 2.7 % in the period 2005–2012. The percentage of Swedish-born women decreased from 87.6 % in 1978 to 83.2 % in 1993 and from 80.7 % in 2005 to 74.9 % in 2012. Even though at present 25 % of women who gave birth were not born in Sweden, the vast majority of these were what would be called “Caucasian” in the USA, and only a small part were likely to be “Black” or “Asian.”

That the mother was not born in Sweden may mean different things. Some were adopted as children and had grown up in Sweden, and the main difference from Swedish-born women would be the genetic setup. Many were immigrants and may have other lifestyle and socioeconomic conditions than Swedish women. There is information on the year of immigration, but it is not routinely included in the Swedish Medical Birth Register.

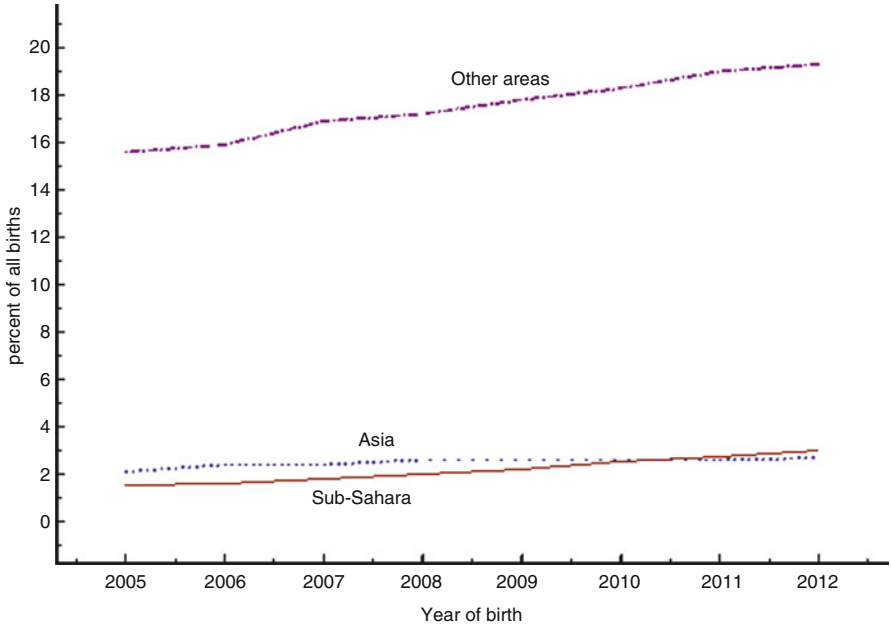


Fig. 8.8 Proportion of women who gave birth and were born outside Sweden, 2005–2012

There are sometimes definite differences in drug use (or registration) in immigrant women compared to Swedish-born women. Some examples can be mentioned taken from Källén (2009). The use of antihypertensive drugs is roughly the same among Swedish-born women and women born in other Nordic countries but clearly lower among other immigrant women. The same is seen for thyroxine and, less clearly, for antibiotics and anticonvulsants. On the other hand, the use of antipsychotic drugs is higher in non-Nordic immigrant women, while antidepressants are used less often than by Swedish-born women. The same is true for anti-asthmatics.

On the other hand the malformation risk is rarely affected by maternal country of birth. Exceptions exist, so for instance the risks for microcephaly and hypospadias are increased and the risk of a cardiovascular malformation or an orofacial cleft is decreased in infants born of immigrant women (Källén 2014).

8.2.9 Socioeconomic Level

For many pregnancy outcomes, socioeconomic variables are of importance, e.g., for preterm birth. The significance of socioeconomic variables varies according to the nature of the society. In some countries, large differences and extreme poverty exist – in other countries like the Scandinavian ones, differences are much less marked, and extreme poverty is rare due to the existing social security system. The significance of socioeconomic level for pregnancy outcome may also depend

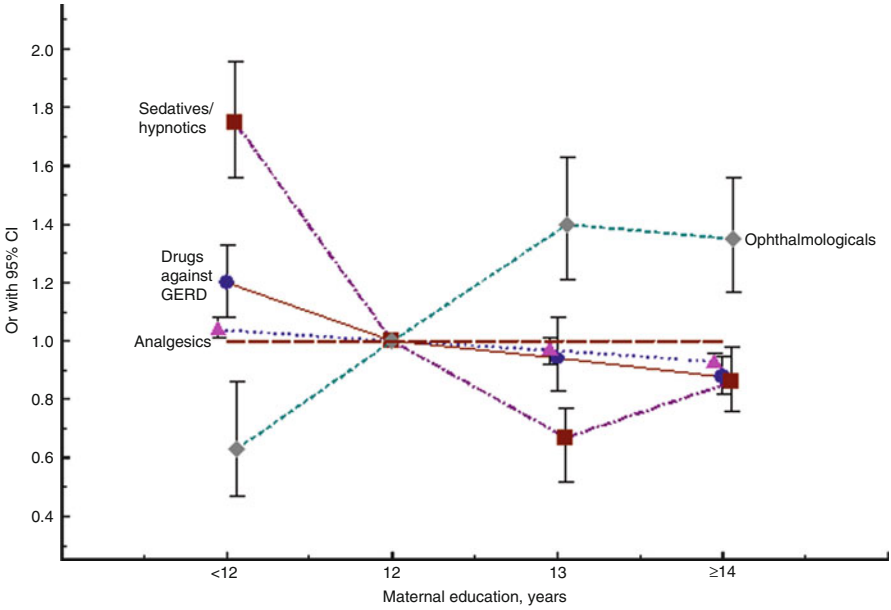


Fig. 8.9 Odds ratio (OR) for the use of some drugs according to length of maternal education

on the health system. In countries where prenatal and delivery care is free and drugs to a large extent are paid by society, the economic aspects are of less importance than in countries where the patients themselves have to pay for care. There may still be socioeconomic differences of significance. In Sweden low socioeconomic level is associated with smoking during pregnancy and obesity, and if adjustment for these factors is made, only a moderate effect of socioeconomy remains. Remaining effects could, for instance, act via nutrition or occupational hazards.

There are different ways to measure socioeconomic level. Sometimes a socioeconomic index has been defined, valid individually or for the area where the woman lives. Such indexes are often built on income, occupation, type of housing, etc. For individual evaluation, the educational level is often used, notably when registers on education exist in the country.

In the Swedish school system, 9 years are compulsory, and most children continue for another 3 years in so-called gymnasium. For many drugs, one can see no variation in usage according to maternal education level, but in some a variation is seen (Fig. 8.9). The use of sedatives and hypnotics is clearly associated with a short education – so is also use of antidepressants (not shown in the Figure). A marked variation in the use of ophthalmologicals is also seen, but then low education level is associated with a low usage and high educational level with a high usage.

The impact of maternal education (adjusted for smoking and BMI) on the risk for congenital malformation is relatively weak (Källén 2014). A classical relationship exists between a low socioeconomic level and infant spina bifida (Elwood et al.

1992), at least in the UK and some parts of the USA. Early studies in Sweden found no relationship between maternal occupation and the risk for neural tube defects in the infants (Ericson et al. 1988).

More evident effects are seen of the parental socioeconomic level on preterm birth, infant birth weight, and intrauterine growth. A large part of this effect is due to smoking and high BMI.

8.2.10 Geography

There may be geographical variations in the recording of drug use, either because of different prevalence of underlying disease, different therapeutic tradition, or different completeness in the ascertainment of drug use. There may also be geographical differences in the registered rates of outcome, for instance, congenital malformations. If both factors are under-ascertained in some areas and completely ascertained in other, it will give a false association between exposure and outcome. The highest risk to get this is if the source of information is the same, for instance, delivery records. If data are obtained from different sources, the risk is reduced but may still exist.

Obviously true differences may also exist in drug usage in different geographical areas, and in true differences in malformation rates and even if the actual risk with the drug is lacking, an apparent effect may appear.

8.2.11 Previous Reproductive History

The previous reproductive history may affect the risk that a new pregnancy results in a malformed infant or some other negative outcome. Such an association may be due to parental genetic or nongenetic factors which increase the risk. To act as confounders, the previous reproductive outcome must also affect the probability of the exposure to drugs.

Previous miscarriages are sometimes associated with an increased risk for a congenital malformation, e.g., a cardiovascular defect, hypospadias, or a multimalformed infant (Källén 2014). For most malformations, the association is weak or absent. The effect is hardly a direct one but more a sign of a common cause, e.g., a chronic condition like preexisting diabetes or a genetic burden, sometimes causing embryonic death, sometimes malformation.

For all three drugs in Fig. 8.10, an increased use is associated with an increased number of previous miscarriages, but the explanation to this may vary.

In case of preexisting diabetes (which is the reason for the administration of insulin), it is likely that the disease increases the risk for a miscarriage so the use of insulin in the present pregnancy indicates a previous risk for miscarriage because most likely the woman was diabetic also during the previous pregnancy.

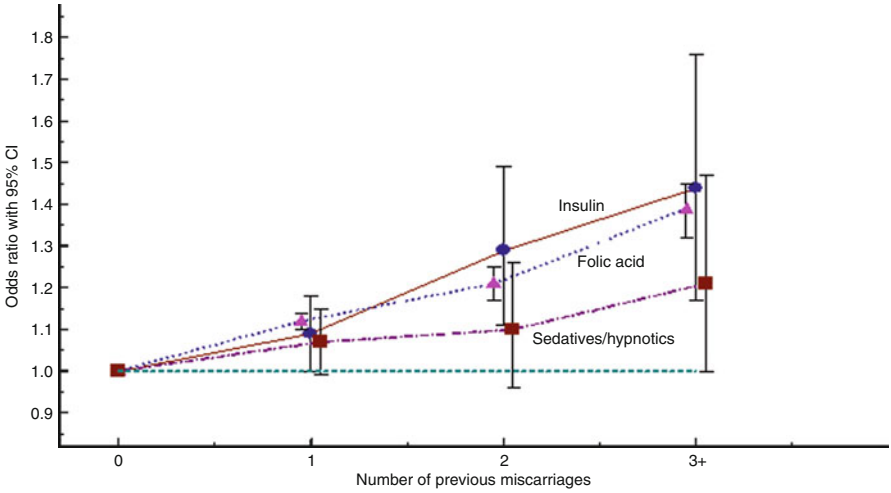
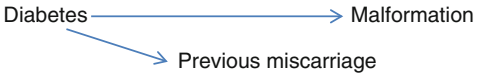
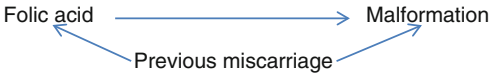


Fig. 8.10 The effect (estimated as odds ratios adjusted for year of birth, maternal age, parity, smoking, and BMI) of the number of previous miscarriages on the use of three drug categories: insulin, folic acid, and sedatives or hypnotics



Folic acid may be taken at an increased rate after previous miscarriages. They will then be a probably weak confounder in an analysis of the effect of folic acid on congenital malformations by increasing the estimate for at least some malformations.



A weak similar confounding may exist for sedatives/hypnotics where previous miscarriages may increase the use of psychoactive drugs, including sedatives/hypnotics. This would result in a slight increase of the malformation risk (which incidentally is very weak for most such drugs). Some careful considerations are needed before adjustment for previous miscarriages is made.

A similar situation exists when an earlier child was malformed as is seen in Fig. 8.11. The presence of a previous infant with malformations is associated with an increased risk for some drugs which indicates the occurrence of chronic diseases with a potential to teratogenesis, e.g., insulin and drugs for chronic hypertension. An increased use of folic acid is also seen (still stronger if the previous malformation was a neural tube defect (OR = 6.0)). It is clinical practice to give high folic acid doses after a previous neural tube defect fetus which explains the strong effect.

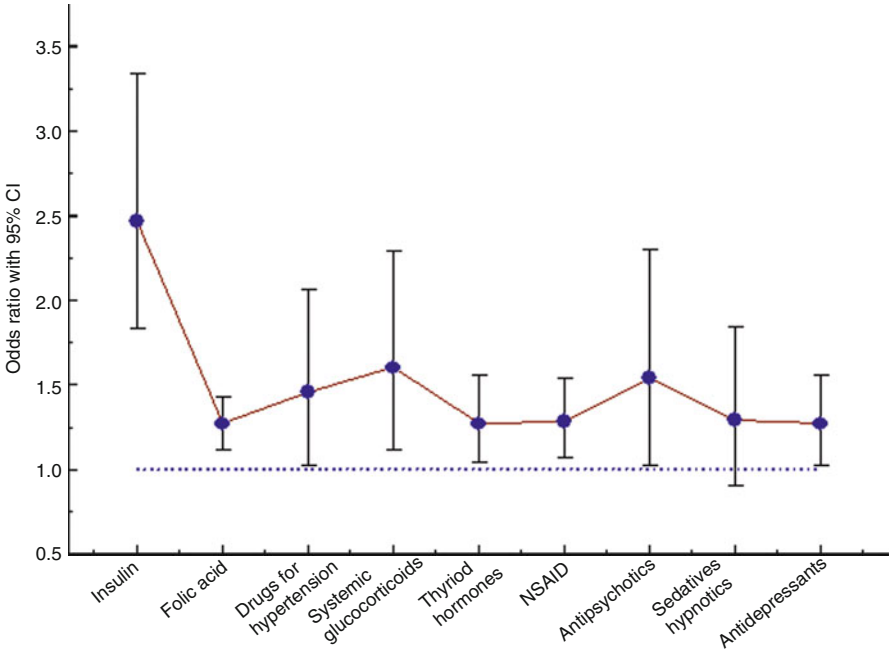
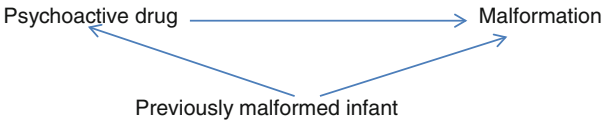


Fig. 8.11 Diagram showing the odds ratios for a number of drug groups, comparing usage among women with and without a previous malformed infant

An increased use of psychoactive drugs can be noted, perhaps from a psychological distress caused by the birth of the damaged infant. In this situation, the birth of a previously malformed infant will be a confounder, notably if a strong heredity for the malformation exists.



There are rather few signs that the occurrence of a malformation in a previous pregnancy would decrease drug use in the following pregnancy. The OR estimate is low for migraine drugs (OR = 0.57, 95 % CI 0.29–1.32), but statistical significance is not reached.

8.2.12 Infant Sex

Many congenital malformations and other outcomes differ in rates between males and females, but infant sex is a confounder only if it also affects the rate of drug exposure. This is rare, but it does happen: nausea and vomiting in pregnancy (NVP) occur more often in pregnancies carrying a female than a male embryo, and if one studies a

malformation which has a deviating sex distribution, infant sex will be a confounder. It is possible that the teratogenic effect of a drug differs on male and female embryos. This can be studied by stratifying for infant sex (but not by adjusting for infant sex).

8.2.13 Concomitant Maternal Disease and Drug Use

Some chronic diseases in the mother can affect infant outcome, including congenital malformations. The most well-known one is preexisting maternal diabetes, but also other diseases may affect embryonic or fetal development, e.g., other endocrine diseases and chronic hypertension. In rare cases also acute diseases may be relevant – a classical example is maternal rubella which carries a high risk for damage of the embryo if occurring in early pregnancy.

Maternal diseases will be relevant for the discussion of confounding by indication (see below) but may also complicate analyses by the effect of other drugs than those used for the disease. So, for examples, diabetes may be associated with chronic hypertension. Usually numbers are not very high but may motivate an exclusion of women with preexisting diabetes in the analysis of the effects of drugs for hypertension.

An expression of comorbidity is the use of other drugs in excess when the effect of a specific drug is studied. This is exemplified in Fig. 8.12.

If drugs used in excess together with the drug under study have a teratogenic effect, this has to be adjusted for or individuals with such exposure should be excluded from the analysis. Obviously such exclusion should then be made also from the non-exposed comparison group. A typical example is the use of anticonvulsants as mode stabilizers with antidepressants. If this is a frequent event, a falsely increased effect of the antidepressant may be obtained. As seen in Fig. 8.12, there is also an excess use, for instance, of drugs for migraines (mainly triptans) but as these have no obvious teratogenic effect they cannot affect the observed effect of antidepressants.

This phenomenon should not be mixed up with the possibility that two drug categories (e.g., antidepressants and sedatives) could act synergistically.

8.2.14 Confounding by Indication

Confounding by indication means that an effect of a drug is really due to underlying disease that was the reason for the drug use. A classic example is diabetes and insulin. There is no difficulty to show that women who took insulin in early pregnancy have an increased risk for having a baby with a congenital malformation. Such infants have a nearly doubled risk to have a major malformation and a three times increased risk of a cardiovascular defect (Källén 2009). The risk for many other types of congenital malformations is also increased. Insulin is reasonably not the cause of this, but the underlying disease, diabetes, is. In order to prove that this is the case one would need to have a large group of pregnant women who got insulin without having diabetes (which is impossible to find) and pregnant women with diabetes type 1 who did not get insulin (which is equally impossible). In practice, insulin is only given to patients with diabetes, and all pregnant type 1 diabetic patients get insulin.

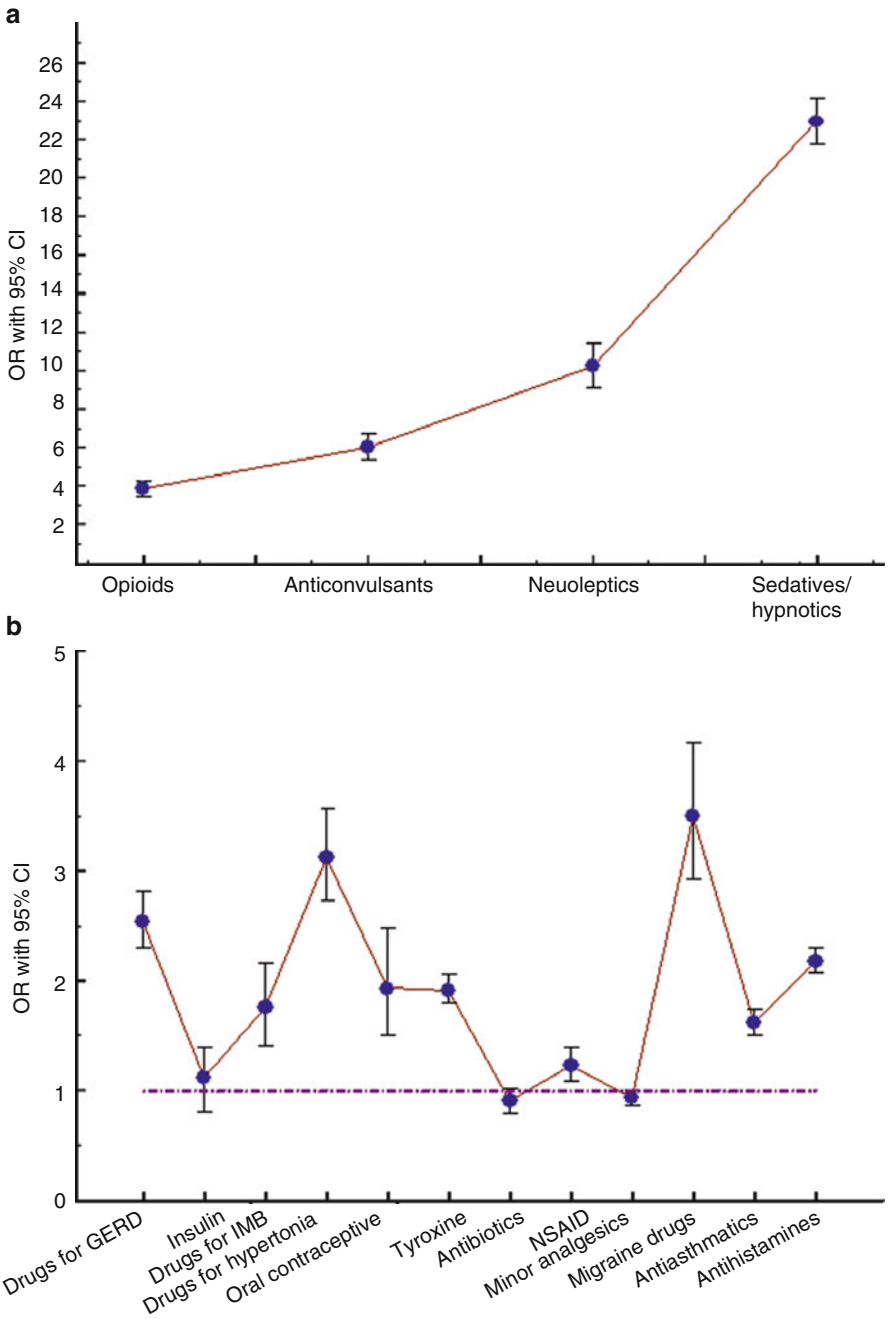


Fig. 8.12 Odds ratio for women having used various groups together with antidepressants versus women who did not use antidepressants, adjusted for year of birth, maternal age, parity, smoking, and BMI. (**a**) shows some drugs with an effect on CNS, (**b**) some other drugs

In other circumstances it is less obvious if the drug or the disease caused the outcome. An example is anticonvulsants. Women with epilepsy who use anticonvulsants generally have an increased risk to get a malformed infant. When this was initially detected, the question came if it was the drug or the disease (notably the occurrence of seizures in early pregnancy) that caused the malformations. It is possible to identify women with epilepsy who did not use anticonvulsant drugs in pregnancy (and they have no increased risk to have a malformed infant), but they reasonably have another or milder form of epilepsy than the women who used anticonvulsants so they are no perfect controls.

Other efforts have been made to separate the effect of drugs from the effect of underlying disease. One has tried to quantify the underlying disease by scoring (propensity scores), often based on disease history before pregnancy, and then in one way or another adjust for disease score. This, for instance, has been made for depression and antidepressants (Oberlander et al. 2006). These authors found an effect on neonatal outcome of the disease but an added effect of the drug use. Other authors found no effect of depression when studying gestational age at birth as outcome (e.g., Suri et al. 2007).

Another way to separate drug effect and disease effect is to compare different drugs, used for the same condition. This can be exemplified with anticonvulsants where the teratogenic effect varies between different drugs – but it is also possible that this is due to the use of different drugs at different forms of epilepsy. In some studies, most SSRI drugs show no signs of teratogenicity, but an association between paroxetine use and cardiovascular defects may exist. Again the use of the specific SSRI drugs may vary between different underlying conditions, e.g., depression and anxiety, so it is still possible that the noted differences between the SSRI drugs are explained by underlying disease.

A further example refers to chronic hypertension and drugs used for that condition. Cooper et al. (2006) found an increased risk for congenital malformations after maternal use of ACE inhibitors but not after the use of other antihypertensives. This would speak for a specific drug effect. A later study of Lennestål et al. (2009) verified the risk increase after ACE inhibitors but found a similar effect also of beta-blockers when used for hypertension. The effect could therefore be a result of the underlying disease, chronic hypertension.

It should be observed that many studies which are quoted as evidence of an effect of underlying disease (e.g., depression and stress) have not distinguished between the disease and the drugs used.

8.3 Interaction Between Confounders

One confounder may interact with another. Above it was shown that the risk for preterm birth varied with maternal age, but that the graphs showing this relationship differed markedly between parity (Fig. 5.1). Furthermore, parity and BMI increase and smoking decreases with maternal age.

Table 8.1 Characteristics of women reporting the use of antidepressants (AD) in early pregnancy and of all women who gave birth, Sweden 2005–2013

Variable	AD, number	AD, percent	Population, number	Population, percent	OR	95 % CI
Year of birth						
2005	881	9.1	91,046	10.3	0.93	0.55–1.02
2006	986	10.1	95,021	10.7	0.96	0.88–1.04
2007	868	8.9	96,299	10.9	0.85	0.78–0.93
2008	1087	11.2	97,594	11.0	1.00	Reference
2009	1171	12.0	99,077	11.2	1.06	0.98–1.16
2010	1274	13.1	104,354	11.8	1.10	1.01–1.19
2011	1188	12.2	99,418	11.2	1.08	0.99–1.17
2012	1318	13.6	99,494	11.2	1.21	1.11–1.30
2013	948	9.8	102,295	11.6	0.84	0.77–0.92
Maternal age						
<20	144	1.5	14,872	1.7	0.74	0.62–0.88
20–24	1131	11.6	115,179	13.0	0.85	0.79–0.91
25–29	2584	26.1	254,911	28.8	1.00	Reference
30–34	3281	33.8	306,239	34.6	1.14	1.08–1.20
35–39	2072	21.3	158,615	17.9	1.40	1.32–1.49
40–44	487	5.0	33,048	3.7	1.52	1.37–1.69
≥45	22	0.2	1736	0.2	1.27	0.82–1.95
Parity						
1	4592	47.2	396,970	44.0	1.00	Reference
2	3060	31.5	323,944	36.6	0.75	0.71–0.78
3	1424	14.6	115,178	13.0	0.86	0.81–0.92
≥4	645	1.6	48,506	5.5	0.80	0.73–0.88
Smoking						
Unknown	74	0.8	45,563	5.2	–	
None	8215	84.5	782,311	88.4	1.00	Reference
<10 cigs/day	1003	10.3	43,695	4.9	2.20	2.12–2.41
≥10 cigs/day	429	4.6	13,029	1.5	3.15	2.86–3.46
BMI						
Unknown	426	4.3	74,930	8.5	–	
<18.5	198	2.0	19,778	2.2	1.00	0.86–1.15
18.5–24.9	4809	50.3	487,427	55.1	1.00	Reference
25–29.9	2528	26.0	203,520	23.0	1.20	1.14–1.26
30–34.9	1056	10.9	69,425	7.8	1.44	1.35–1.54
≥35	624	6.4	29,518	3.3	1.94	1.78–2.11
Total number	9721	–	884,598	–	–	–

Odds ratios (OR) with 95 % confidence intervals (95 % CI) adjusted for all other variables in the table

In order to study the impact of a specific confounder, adjustment for other confounders should be made. Often tables are produced which show the distribution of for instance maternal age and parity – it is more informative to calculate the odds ratio for each maternal age adjusting for parity and for each parity adjusting for maternal age.

Table 8.1 shows the two ways to tabulate characteristics of women using antidepressants compared with other women. The adjusted ORs thus show the effect of each variable irrespective of the other variables in the Table. This makes it possible, for instance, to evaluate if the maternal age affects drug use irrespective of parity and if parity does so irrespective of age.

8.4 Residual Confounding

Even when extensive efforts have been made to adjust for confounding, residual confounding may remain. Either important confounders exist which have not been identified or it has not been possible to include them in the adjustment, or adjustment has been made for a factor but in an incomplete way. If adjustment for maternal smoking has been based on three levels: none, <10 cigarettes per day, and ≥10 cigarettes per day, no complete correction has been made for smoking 20 cigarettes per day. If in a group of women more smoke ≥10 cigarettes per day than another group, it is likely that the proportion smoking ≥20 cigarettes a day are also overrepresented and that has not been possible to adjust for. If adjustment for the two known levels of smoking results in a reduction of the risk estimate, it is likely that a further reduction had been obtained if adjustment could have been made for ≥20 cigarettes per day.

In nearly every study, it is possible that residual confounding exists. The larger effect one has obtained by the adjustments applied, the more likely it is that the association is sensitive for residual confounding.

References

Blomberg M, Källén B (2010) Maternal obesity and morbid obesity: the risk for birth defects in the offspring. Birth Def Res (Part A) 88:35–40

Cedergren M (2006) Effects of gestational weight gain and body mass index on delivery outcome in Sweden. Int J Gyneacol Obstet 93:269–274

Cooper WO, Hernandez-Dias S, Arbogast PG, Dudley JA, Gideon PS, Hall K et al (2006) Major congenital malformations after first-trimester exposure to ACE-inhibitors. N Engl J Med 354:2443–2451

Elwood JM, Little J, Elwood JH (1992) Epidemiology and control of neural tube defects. Monographs in epidemiology and biostatistics, vol 20. Oxford University Press, Oxford, New York, Tokyo

Ericson A, Källén B, Löfqvist E (1988) Environmental factors in the etiology of neural tube defects. A negative study. Enviroment Res 45:38–47

Gnoth C, Godehardt E, Frank-Herrmann P, Friol K, Tigges J, Freundl G (2005) Definition and prevalence of subfertility and infertility. Hum Reprod 20:1144–1147

Hackshaw A, Rodeck C, Boniface S (2011) Maternal smoking in pregnancy and birth defects: a systematic review based on 173 687 malformed cases and 11.7 million controls. Hum Reprod Update 17:589–604

Sundh KJ, Henningsen A-J, Källén K, Bergh C, Romundstad LB, Gissler M et al (2014) Cancer in children and young adults born after assisted reproductive technology: a Nordic cohort study from the Committee of Nordic ART and Safety (CoNARTaS). Hum Reprod 29:2050–2057

Jones KL, Smith DW, Streissguth AP, Myrianthopoulus N (1973) Outcome in offspring of chronic alcoholic women. Lancet i:1076–1078

Källén B (2009) Drugs during pregnancy. Nova Biomedical Books, New York

Källén B (2014) Epidemiology of human congenital malformations. Springer, Heidelberg, New York, Dordrecht, London

Källén B, Finnström O, Lindam A, Nilsson E, Nygren K-G, Otterblad OP (2010) Cancer risk in children and young adults conceived by in vitro fertilization. Pediatrics 126:e270–e276. doi:10.1542/peds.2009-3225

Källén K (1998) Förlossningsresultat bland invandrarkvinnor i Sverige. (Delivery results in immigrant women in Sweden). EpC Report 1998:1. Socialstyrelsen

Källén K (2001) The impact of maternal smoking during pregnancy on delivery outcome. Eur J Pub Health 11:329–335

Källén K (2002) Maternal smoking and congenital malformations. Rev Fetal Maternal Rev 13:63–86

Lennestål R, Otterblad Olausson P, Källén B (2009) Maternal use of antihypertensive drugs in early pregnancy and delivery outcome, notably the presence of congenital heart defects in the infants. Eur J Clin Pharmacol 65:615–625

Mission JF, Marshall NE, Caughey AB (2015) Pregnancy risk associated with obesity. Obst Gynecol Clin North Am 48:335–353

Oberlander TF, Warburton W, Misri S, Aghajanian J, Herzman C (2006) Neonatal outcome after prenatal exposure to selective serotonin reuptake inhibitor antidepressants and maternal depression using population-based linked health data. Arch Gen Psychiatry 63:898–906

Oberlander TF, Warburton W, Misri S, Riggs W, Aghajanian J, Herzman C (2008) Major congenital malformations following prenatal exposure to serotonin reuptake inhibitors and benzodiazepines using population-based health data. Birth Def Res (Part B) 83:68–76

Reis M, Källén B (2013) Combined used of selective serotonin reuptake inhibitors and sedatives/hypnotics during pregnancy; risk of relatively severe congenital malformations or cardiac defects. A register study. BMJ Open 3:e002166. doi:10.1136/bmjopen-2012-002166

Smith LK, Draper ES, Evans TA, Field DI, Johnson SJ, Manktelow BN et al (2015) Associations between late and moderately late preterm birth and smoking, alcohol, drug use and diet: a population-based case-cohort study. Arch Dis Child Fetal Neonatal Ed 0:F1–F6. doi: 10.1136/archdischild-2014-307265

Suri R, Altschuler L, Hellenmann G, Burt VK, Aquino A, Mintz J (2007) Effects of antenatal depression and antidepressant treatment on gestational age at birth and risk of preterm birth. Am J Psychiatry 164:1206–1211

Statistics for Dummies

9

9.1 Risk Estimates

In order to epidemiologically investigate if an exposure (e.g., maternal use of a drug) causes an outcome (e.g., infant congenital malformations), one first has to look for an association between the two; if the outcome occurs more often than expected after the exposure. If so, the next step is to see if it is likely that the excess probably is random or if it is "statistically significant."

The crucial point is then to determine the number of outcomes after the exposure and to get an estimate of how many one would have had if the exposure had not affected the outcome, that is, the *expected* number. You sometimes hear that you don't expect a baby to be malformed (which is true), but "expected" here means what number you would expect if the exposure had no effect. The expected number has to be estimated from some sort of a control material, for instance, from the occurrence of the malformation in non-exposed infants.

If the observed number of outcomes exceeds the expected number, this indicates the presence of a risk associated with the exposure. In order to quantify this risk, one can use a *risk ratio*, that is, the ratio between the risk after exposure divided with the risk after non-exposure. We can go back to the simple 2 × 2 table shown already in the introduction. It consists of four central cells and the rand sums (totals).

	Outcome	No outcome	Total
Exposed	N_1	N_2	N_1+N_2
Unexposed	N_3	N_4	N_3+N_4
Total	N_1+N_3	N_2+N_4	$N_1+N_2+N_3+N_4$

The risk after exposure is thus $N_1/(N_1+N_2)$ and after non-exposure $N_3/(N_3+N_4)$ where N's represent numbers in each cell and the risk ratio will be $N_1/(N_1+N_2)$ divided by $N_3/(N_3+N_4)$. If the risk ratio = 1, the risks are identical, and the exposure has no effect – if it is over 1, the exposure may increase the risk of outcome; if it is

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_9

lower than 1, the exposure may protect against the outcome. Note thus that at null risk, the risk ratio estimate is 1.0, not 0.

The expected number in the "exposed, outcome" cell will be a product of the number of exposed infants and the risk in the total material, thus $(N_1+N_2)*(N_1+N_3)/(N_1+N_2+N_3+N_4)$.

Such risk estimates are difficult or impossible to calculate from case-control studies as the proportion between outcome and no outcome is already defined, for instance, two controls for each case.

The odds for exposure in the outcome is N_1/N_3 and in the non-outcome N_2/N_4 and the odds ratio is $(N_1/N_3)/(N_2/N_4)$. Note that this will be the same if one instead divides the odds for outcome in the exposed with the odds for outcome in the non-exposed.

The odds ratio will with necessity always be larger than the risk ratio (because you divide with a smaller number), but when outcomes are relatively rare (like congenital malformations) and exposures are relatively rare (like most drug use), the difference between the odds ratio and the risk ratio will be small.

In cohort studies one can calculate both odds and risk ratios but in order to get a similar measure as that used in case-control studies, often odds ratios are used even though risk ratios would be more logical.

A risk ratio of 1.7 thus means that after exposure, the outcome is 70 % more common than without exposure. One should not mix this up with the absolute risk – the risk a woman who has taken a drug in early pregnancy has to get a malformed baby, for instance. If we say that the general risk to have a baby with a major malformation is 3 %, the 70 % risk increase means that the risk increases to 5.1 %. If we are talking about a spina bifida, where the risk in an unexposed fetus is perhaps 1/2000 (=0.0005), the 70 % risk increase will mean an absolute risk of 1/1176 (=0.0005*1.7). Moderate risk increases are usually of limited interest for the individual, but if the exposure is common in the population, they can cause a substantial number of damaged outcomes.

9.2 Is the Odds Ratio or Risk Ratio Statistically Significant?

Let us start with a simple situation where no adjustments are made, but we have a simple 2 × 2 table as shown above. We can take a hypothetical example of maternal use of a drug and the risk to have an infant with a cardiovascular defect. We have collected data for 1000 drug users and compared them with 1000 nondrug users. Among the latter, 1 % had a cardiovascular defect; among the former, 2 % had such a defect – the risk ratio is thus 2.0. The distribution of the numbers is seen in the table below.

	Malformation	No malformation	Total
Drug use	20	980	1000
No drug use	10	990	1000
Total	30	1970	2000

The odds ratio is thus 20/980 divided by 10/990 = 2.02, very close to the risk ratio of 2.0. But is the difference certain or could it be random? To answer this, one can apply a chi-square test. If this is done, you will find that the chi-square is 3.38 which corresponds to a probability (p-value) of 0.07. If we have decided that we accept statistical significance if $p < 0.05$, the answer is thus: no, this can be a chance finding.

The chi-square test compares the observed and expected numbers in each of the four central cells in the 2×2 table and is based on a squared normal distribution.

In this calculation we have made the assumption that the cardiac defect rate after drug use could be either higher or lower than after no drug use, what is called a two-sided test. If we are absolute sure that the use of the drug cannot protect against getting an infant with a cardiovascular defect, then one can apply a one-sided test and then $p = 0.03$, thus statistically significant. It is customary to always use two-sided tests, but it is sometimes useful to consider what happens if you apply a one-sided test. There are actually situations when the use of a drug is associated with a lower than expected malformation risk, e.g., use of antihistamines at NVP – not because the drug protects against malformation but because NVP is associated with a well-functioning placenta and therefore with a slightly decreased risk for malformation. Therefore, drugs used for NVP may actually show a reduced risk.

The chi-square test is built on normal distributions, and when numbers are low, such approximations are not allowed. One can then use an exact test, Fisher test. If one applies this on the example above, the exact two-sided p-value is 0.10, thus slightly weaker than that calculated with the chi-square test (0.07). If we instead had 4 malformed infants in the exposed group and 2 in the non-exposed group (which would give the same risk ratio), the chi-square test would give $p = 0.41$ and the Fisher test $p = 0.68$ – with so few cases, only the Fisher test should be used. As a rule of thumb, when the expected number in the smallest cell is less than 10, chi-square tests should be replaced by exact tests.

The importance of the p-value should be looked upon with some light-heartedness. A significant p-value does not necessarily prove that the association is true. A p-value of 0.04 (which is thus significant according to the most commonly used definition) only means that the chance that the finding is random is 1 in 25, comparable to the chance to draw a red king from a pack of cards on the first trial, which of course may happen and does not prove that you are an unusually clever finder of a red king. An association which is not statistically significant may well be true but the data so far do not show that they are, *and it does not prove that no association exists*. The difference between a p-value = 0.049 and one = 0.051 is not very large, but the former is regarded as significant, but not the latter. We will come back to this later in the text.

This may also be the place to point out the difference between a statistically significant effect and a clinically significant effect. The latter has two aspects: the significance for the individual case and the significance for the occurrence of the outcome in the population. If there is a moderate increase (say, a 50 % increase) of a common malformation (say a cleft lip/palate), this is of little importance for the

individual which has been exposed. If the unexposed infant has a risk of such a defect amounting to 1/1000, it will increase after exposure to 1/670. If, however, 670 women will use the drug in question, it will result in one extra cleft case. A good example of this is the effect of maternal smoking on malformation risk – a rather low-risk increase which was of importance in the population because so many women smoked.

9.3 The Confidence Interval

The confidence interval of a risk or odds ratio estimate indicates within which range the true ratio lies with, for instance, 95 % certainty – the 95 % confidence interval (95 % CI). This is actually more informative than a *p*-value. If we take the example in the table above and apply a Fisher exact test, we find the 95 % CI of the odds ratio to be 0.90–4.86. This interval thus tell us that the true odds ratio may be so low as 0.90 (indicating a protective effect of the drug) or as high as nearly five times increased. Obviously it is not very likely that the true odds ratio is nearly five, but it may be so. The fact that the no-effect odds ratio (1.0) falls within the confidence interval tells the same as the *p*-value: the registered increase can be random.

If we instead had found 30 cases in the exposed group the odds ratio had been 3.1 (the risk ratio 3.0), the *p*-value 0.001 and the 95 % CI of the odds ratio would be 1.45–7.06. The *p*-value tells that it is only one chance in 1000 that the finding is random, and the confidence interval shows that the odds ratio is at least 1.5 and may be as high as 7.1.

In the table above, we compared two equally large groups, each consisting of 1000 individuals. If we instead had a control material of 10,000 individuals (thus with 100 infants with heart defects), the same odds ratio would exist, but the *p*-value would be 0.009 and the 95 % confidence interval 1.18–3.31, thus much narrower than in the example above. The reason is of course that the large control material can estimate the expected number of malformations among the 1000 exposed infants with much higher precision that what the smaller control material could.

9.4 Expected Numbers

We have talked about the expected numbers with which the observed numbers should be compared. In the standard 2 × 2 table as exemplified above, there are four central cells where observed and expected numbers should be compared. The expected numbers are calculated from the totals in the table – if there is no effect of the drug use, the proportions of malformed and non-malformed infants should be the same in both rows. We can calculate the expected numbers for each one of the four cells:

Exposed malformed: 1000*30/2000 = 15
Non-exposed malformed: 1000*30/2000 = 15

Exposed non-malformed: 1000*1970/2000 = 985
Non-exposed malformed: 1000*1970/2000 = 985

The numbers will be pairwise the same because there are equally many exposed as non-exposed individuals.

A chi-square calculation adds the values of (observed-expected)^2/expected for the four cells: $(20–15)^2/15 + (10–15)^2/15 + (980–985)^2/985 + (990–985)^2/985 = 1.67 + 1.67 + 0.025 + 0.025 = 3.38$. For getting a p-value =0.05, one needs a chi-square = 3.85. As can be seen in this calculation, the major part of the chi-square values comes from the two small cells (the malformation cells).

When checking the p-value for a calculated chi-square result, one meets the concept of degrees of freedom (d.f.). If one value is changed in a 2×2 table, the other three values are also changed. Therefore, it has only one d.f. If one has a table with more cells, say, n vertical and n1 horizontal, the d.f. will be $(n - 1)*(n1 - 1)$.

The table shown above illustrates a so-called hypergeometric distribution. If the non-exposed material is very large (for instance, if one compares a group of women who used a specific drug with all women who gave birth in the population), the uncertainty in the distribution will be nearly exclusively located to the exposed group, and one could look upon the rate of malformations in the non-exposed group as relatively exact. Then only the two outcomes remain (malformed and non-malformed infants in the exposed group), and they will distribute as in a binomial distribution. One can go one step further if the malformed group is small compared with the non-malformed group. Then the binomial distribution can be approximated as a Poisson distribution. This means that we can get a good idea just by comparing the observed number of malformed infants in the exposed group with the expected number. The beauty is that one can then evaluate also small numbers using exact Poisson distributions. If we, for instance, have five infants with a rare malformation, one can from the Poisson distribution learn that the 95 % confidence interval of 5 is 1.62–11.7, and if the expected value is less than 1.62, it is likely that the increase in malformation rate is not random.

9.5 Dealing with Confounders

So far we have not bothered about confounders, which will now be discussed. Adjustment for confounders is supposed to eliminate the effect of confounding. If, for instance, we want to adjust for maternal age, we compare exposed and non-exposed infants with consideration to possible differences in maternal age distributions in the two groups. This can be done at different levels of precision, from adjustment for 1-year maternal age, via adjustment for 5-year maternal age groups to crude adjustments, e.g., <30 and ≥30 years. The adjustment will of course be more complete if the adjustment refers to 1-year or 5-year age groups than if it is just two crude groups. Usually, a 5-year adjustment is enough precise, but in certain

circumstances, one should make more detailed adjustment, e.g., for 1-year groups above 35 in studies of Down syndrome or below 25 years in studies of gastroschisis.

9.5.1 Matching

In a case-control study or when two cohorts are compared where one is non-exposed, the control or non-exposed group of individuals can be chosen in a way to resemble the cases or the exposed cohort. This is made by a matching procedure. One possibility is to select one or more controls to each case with the same confounding characteristics as the case, say born the same year, with same maternal age, parity, smoking habits, and BMI. In this way one gets controls which are specifically selected according to case characteristics. They may make up pairs (case, control) or triplets (case, two controls) or sets with more than two controls per case. The most efficient mode of analysis consists of comparisons within such pairs, triplets, etc. If we take the simple example of one case-one control and are interested in an exposure which either is there or is lacking (e.g., maternal drug use), you will get four types of pairs:

1. Case exposed, control exposed – n1 such pairs
2. Case non-exposed, control non-exposed – n2 such pairs
3. Case exposed, control non-exposed – n3 such pairs
4. Case non-exposed, control exposed – n4 such pairs

Among these pairs, only n3 and n4 are informative. If the exposure does not increase the risk for the outcome (congenital malformation), n3 and n4 will be about the same. If the exposure increases the risk for a malformation, $n3 > n4$. To evaluate how likely a difference is true or not, one applies a binomial distribution test. If, for example, in a study we find $n3 = 120$ and $n4 = 80$, we can look up what 95 % CI this binomial distribution has (120/200) – the distribution is 60 % exposed, and this has a 95 % CI of 53–67 % and is thus higher than the 50 % which would be the case in an absence of the effect. If we had only a fourth as many (30/50), the distribution would still be 60 % exposed, but the 95 % CI would be 45–74 %, and the 50 % corresponding to no effect of exposure lies within the confidence interval so the effect may be random.

When we have triplets instead of pairs, the calculations get more complex but can be made in a corresponding way.

Another way to match controls to cases is by so-called frequency matching. Then a group of controls is selected with the same distribution of the confounding factors as the case group has, for instance, a similar maternal age, parity, and smoking distribution. In that situation, group comparisons are made.

Matching has to be made before the data collection process and will be firmly rooted. If new confounding features of interest appear, it is too late to include them in the matching, but adjustments must be made.

One sometimes finds tables comparing the case and control group for matching criteria. To add statistical tests studying if the two groups differ from these points of view is of course nonsense. Statistical tests should decide if differences can be caused by chance – and here we know that the two groups are similar because of the matching.

9.5.2 Adjustment

There are two main methods for adjusting, the logistic regression and the Mantel-Haenszel test (Mantel and Haenszel 1963). One can look upon the latter method as consisting of a series or strata of 2 × 2 tables, one for each situation of confounding, e.g., one table valid for mothers aged 25–29 years, having their first baby, nonsmoking, and with a BMI of 25–29. In an analysis of Table 8.1, it would be a total of 6048 such tables. The method gives a chi-square value based on one d.f. and estimates the average relationship between the exposure and the outcome. It may vary between different strata which can be controlled by separate analyses of, for instance, women above and below 30 years age. From the chi-square, the confidence interval can be calculated, e.g., with the simple but approximate method of Miettinen (1974).

The main problem with the Mantel-Haenszel technique is that there must be data from the non-exposed individuals in every 2 × 2 stratum with exposed individuals. If this is not the case, the stratum cannot be used and information is lost. When we are dealing with very large control groups (like using all infants in the population), the risk for this to happen is small. When smaller data sets are analyzed, this can be a major problem.

Nowadays one mainly uses a logistic regression model for adjustment of confounders. In such an analysis all data can be used, because the control value for each case is estimated from a regression which usually is linear but could be polynomial. A correct use of a logistic regression method necessitates a well-modeled regression which may sometimes be difficult to construct. If we look at the graph showing the relationship between maternal age and risk of preterm birth (Fig. 5.1), it is obvious that a straight line cannot correctly describe the relationship and that the relationship varies between different parities.

In the standard model, the basic formula looks like the following:

If p is the rate of occurrence of a specific event in the material and one wants to adjust for n different variables, then

$$\ln\left(p/(1-p)\right) = \alpha + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \cdots + \beta_n * X_n.$$

β_1 to β_n are thus regression coefficients and X_1 to X_n data for the n different variables. X_1 could, for example, be drug use (0/1) and X_2 maternal age. Using an iterative technique, the best fit of the equation to the data can be obtained, and one also estimates (with errors) the coefficients of each term (independent of the effects of the other terms). The regression coefficients can be transformed into ORs.

It is important to realize that the result tells the effect of each variable independent of the effects of the others, that is, it is adjusted for the other variables. If two added variables are included which have a strong relationship, they will more or less kill each other. Entering both pregnancy duration and birth weight will result in information on birth weight at each gestational week, that is, the result of intrauterine growth. If one wants to know if one variable has different effects according to the presence of another variable (e.g., smoking and obesity), one can introduce an interaction term (X_1*X_2) – if significant it means that both variables have an effect and that the size of each depends on the other variable.

9.6 Survival Analysis

This is especially valid in studies of long-term effects, e.g., childhood survival or development of chronic diseases. The basic method is the Kaplan-Meier test which describes number of events (e.g., death or diagnosis of a chronic disease) for each time period (e.g., year) among number of individuals "at risk." This number will gradually decrease, partly because individuals die or get the disease (and therefore cannot get the disease again), partly by the fact that some individuals are lost for follow-up (censoring), perhaps because of refusal of participation or emigration. The method gives one survival graph for each exposure situation, e.g., maternal use or non-use of a drug.

The Cox regression method also follows "survival" but makes it possible to add various confounding variables and will in this way be similar to the logistic regression method and usually necessitates linearity and proportionality.

A common method to study long-time effects is to calculate the number of events of a specific disease (e.g., first ADHD diagnosis) per number of follow-up years. This needs proportionality in the data. If we study 200 newborn individuals for 10 years, we get 2000 years of observations, but the same number is obtained if we study 2000 individuals for 1 year. In the former group, a number of individuals will develop ADHD, in the latter group probably none will get that diagnosis because one seldom gets a diagnosis before 1 year's age.

9.7 Power Analysis

A power analysis should be made when a study is planned and should answer the question: how many individuals do I have to include in the study to be able to demonstrate a certain risk increase? Or what size of a risk do I have a chance to detect given the number of individuals I can study?

If we take the question if maternal use of a drug causes an increased risk for any major malformation, we need to know the rough prevalence of malformations in the study population. We also need to know the design of the study and the number of controls or non-exposed individuals per case we can put up.

We will make two assumptions that we want an 80 % chance to detect an association ($\beta = 0.80$) with statistical significance ($\alpha = 0.05$). These two values can of course be changed.

Let us take an example. We will study two cohorts, one with maternal use of the drug and the other without such use. We believe that the population risk of any major congenital malformation is 3 %, and we want to be relatively sure to identify a doubling of the risk. If we have one unexposed individual per exposed individual, we would need 749 exposed and 749 unexposed individuals. By increasing the unexposed group to 16,300 (which is possible if exposures make up 2 % of all, e.g., antidepressants), we could reduce the number of exposed to 326 – further increase will only marginally decrease that number. This number of exposed individuals should – with 6 % malformations – result in 19 infants with major malformations.

If we instead plan a case-control study with two controls per case, and the exposure rate among controls is 2 %, we would need 875 cases and 1750 controls.

The best chance here is obviously to study outcome in exposed and unexposed pregnancies, where unexposed are represented by all other pregnancies in the population. To detect a doubling of the rate of cardiovascular defects, one would need 980 exposed individuals, the corresponding figure for orofacial clefts would be 2205, for spina bifida 40,822, and for gastroschisis 68,046.

Let us take the second problem: how high-risk increase can we detect given different numbers of exposures. Figure 9.1 shows the power to detect an increase in

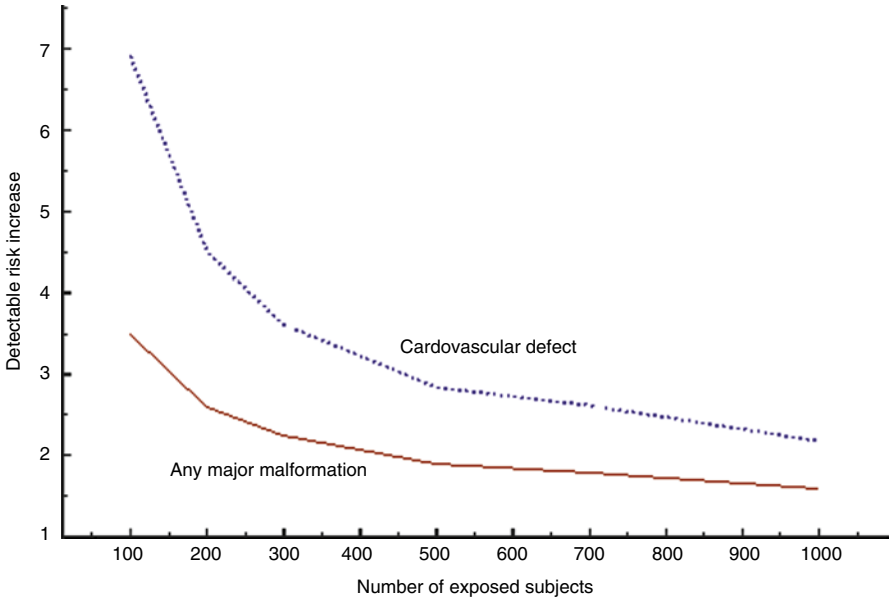


Fig. 9.1 Diagram showing the risk increase detectable at different numbers of exposed subjects for any major congenital malformation (3 % in population) or cardiovascular defects (1 % in population). The exposure rate in the population is supposed to be 2 %

any major malformation and in any cardiovascular defect at different numbers of exposed subjects and 50 times as many unexposed subjects (corresponding to 2 % exposure risk), supposing $\alpha = 0.05$ and $\beta = 0.80$. With 1000 exposures, one would need an 11 times increased risk of spina bifida to be detectable if the background rate is 1/4000 – seven times increased risk if the background rate 1/2000.

These calculations underline the futility in small studies when they refer to congenital malformations. Only extremely high risks (like that caused by thalidomide) can be detected, and negative studies have very little information value.

9.8 The *p*-value and Mass Significance

Researchers have a tendency to be blinded by *p*-values. Here are some common misunderstandings:

A statistically significant p-value proves that the exposure causes the outcome. First of all it only suggests that the observed association may not be caused by chance – it may be, however, because every 20th time one make a test, the *p*-value may randomly lie under 0.05 (if that is the significance level, one has determined). These 20 times can be 20 tests made on the same material (looking, for instance, on five different drugs and four different outcomes). This can be corrected for by statistical methods. It could also be one of 20 studies on the same problem, made in different parts of the world, published or unpublished.

Second, the *p*-value only indicates that there most likely is an association between exposure and outcome, but it may not be causal, the exposure may not have caused the outcome but the association is due to confounding (see above). Sometimes one has to consider the possibility that the outcome has caused the exposure, so-called reversed causality.

Absence of a statistical significance means that the exposure does not cause the outcome. The correct interpretation is that the study is not large enough (does not have enough power) to show that the observed risk difference is not caused by chance – it may well be true anyway. A *p*-value of 0.053 or a lower confidence limit of 0.98 is, strictly speaking, expressions of nonsignificance – but common sense should regard such findings as at least suggestive. If the chance for randomness is 1/20 or 1/19 is not that important.

If one group is statistically significant and, the other is not, the two groups differ. This is no proof that they differ – this has to be shown by an analysis which demonstrates that the risk estimates for the two groups cannot be estimates of the same risk.

Mass significance is an expression of multiple testing that one does not restrict the analysis to one predetermined association but make tests on multiple situations. If we are in the situation that we have a population study where any kind of drug use

has been registered and any type of congenital malformation can be studied, we will have the possibility to perform an enormous number of tests – and some of them will come out "statistically significant" just by chance. If we have studied 100 different drugs and 20 different malformations, we can make 2000 tests, and there should randomly be 100 "significant" deviations, either as risks or as protective effects. We will later in this book discuss ways to find out which of them are likely to be true and which are likely random – but it is not possible to do it from the original material.

There are more refined methods of mis-use of mass significance. One way is to start the study mentioned above with 100 drugs and 20 malformations, look for apparent risk increases, and select them for statistical testing, not mentioning the other possible tests which could have been made. The *p*-value has a meaning only if the test was decided before data were available – one has put this as "a *p*-value must have a history." If you decide to make a study of drug A and malformation B and collect data on all drug use and all malformations, you can select the A-B association for testing and the *p*-value has a meaning. If you did not select the A-B association a priori but did it because it seemed to be interesting, the *p*-value has no meaning.

Another method to get what you want is to produce a set of different groups and then compare the group with the highest value with the group with the lowest value, again defining such groups from the outcome. If many exposure groups are formed, one has to decide before data are collected which groups to compare.

References

Mantel N, Haenszel W (1963) Statistical aspects of data from retrospective studies of disease. J Natl Cancer Inst 32:719–748

Miettinen OS (1974) Simple interval estimation of risk ratio. Am J Epidemiol 100:515–516

Lumping or Splitting? 10

10.1 Lumping or Splitting of Drug Exposures

This can be done on different levels. An example is shown in Table 10.1

The large group of any antidepressant or the main subgroups, TCA and SSRI, thus do not show any significant risk increase – but the specific antidepressants clomipramine and paroxetine do. If this effect is true, it is hidden in the analysis of larger groups as it is based on only 6 and 7 %, respectively, of the total material. It may, however, be false because when we divide the main group into a number of subgroups, a risk for a "multiple testing" effect exists. The paroxetine association has been seen in some but not all studies performed but is supported by a recent meta-analysis (Bérard et al. 2016).

Sometimes one can lump drugs of quite different nature but with some effect or side effect in common. Thus, for instance, some studies concentrated on folic acid antagonists (Hernández-Diaz et al. 2000, 2001; Matok et al. 2009), others on nitrosatable drugs (Olshan et al. 1989; Gardner et al. 1998). Other similar groupings can be made. A number of drugs with quite different pharmacological effects may cause QT interval prolongation, a mechanism which in experimental systems has been suggested to be teratogenic (Danielsson et al. 2013). A problem in such studies is the inclusion of drugs with a known teratogenic effect, e.g., many anticonvulsants. If they make up a large part of the studied group, they may result in a significance which may be unrelated to the characteristic after which the studied group was selected.

10.2 Lumping or Splitting of Outcomes

Typically this problem arises when congenital malformations are studied. The concept of congenital malformations is so wide and contains conditions with different etiology and embryology. A teratogenic drug could affect a central component in embryonic development like cell division, resulting in widespread malformations, or has a more specific effect like an antiandrogenic effect which could affect male genital development and cause hypospadias.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_10

Table 10.1 Association between maternal use of antidepressants and infant cardiovascular defects

Drug group	Number of exposures	Number of cardiovascular defects	OR	95 % CI
Any antidepressant	23,658	208	1.04	0.91–1.20
TCA	2139	36	1.29	0.93–1.79
Clomipramine	1399	28	***1.49***	***1.03–2.15***
SSRI	19,181	151	0.94	0.80–1.11
Paroxetine	1687	24	***1.67***	***1.12–2.50***

From Källén et al. (2013)
Bold text marks statistical significance
TCA tricyclic antidepressant, *SSRI* selective serotonin reuptake inhibitor

It is reasonable to study all congenital malformations together. This is actually what the pregnant woman is interested in: does the use of the specific drug she has taken increase the risk for a congenital malformation? She might restrict this to major or serious malformations – she may not be concerned about minor or easily corrected anomalies – but it is still a very heterogeneous concept. To get an effect on the total risk of significance for an individual patient, the drug must affect a common malformation or a group of malformations, e.g., congenital heart defects which make up a substantial part of all major malformations. If the drug selectively doubles the risk for spina bifida, the absolute risk is still very low and an increase in total malformation risk is not detectable.

There is thus a need to group malformations into more homogenous groups in order to identify effects on specific embryonic processes. This has very little to do with what organ the malformation belongs to, the subdivision principle used by the ICD code. This was discussed earlier in this book (Chap. 6). The "lumping" should be made with some consideration of the pathogenetic pathways of different malformations. An example is, for instance, to group neural tube defects (anencephaly, spina bifida, encephalocele) – even though some data actually indicate that anencephaly, encephalocele, and upper spina bifida represent one entity and lower spina bifida another. The former are defects of the closure of the neural tube, the latter origin in the most caudal part of the neural rudiment which develops from the so-called caudal eminence.

Another example is that both esophageal atresia and anal atresia develop by a similar process at about the same time and occur together more often than what chance can explain. Some cases of small gut atresia are probably also related even though other cases may have a quite different pathogenesis.

In order to identify possible specific patterns of malformations among infants exposed to a drug, it is a good idea to list the observed malformations – only in large studies of common drugs will the numbers of malformed infants be too large to permit a detailed listing. If such a listing is made and one notices that a number of cases seem to be identical or resemble each other from a pathogenetic point of view, the statistical evaluation of the observed "cluster" is uncertain. If there was no prior hypothesis to explain the "cluster," it is not reasonable to test the observed number

Table 10.2 Data on malformations observed among 151 infants exposed in utero for methimazole

Malformation	Number	Note
Lacrimal duct stenosis	1	Minor
Ventricular septum defect	1	Major
Single umbilical artery	1	Minor
Unspecified cardiac defect	1	Major
Choanal atresia[a]	1	Major
Choanal atresia[a] + hypospadias	1	Major
Esophageal atresia[a] + ventricular septum defect + aorta malformation	1	Major
Ileum atresia	1	Major
Meckel diverticle[a]	1	Minor
Omphalocele[a]	1	Major
Gastroschisis	1	Major
Hydronephrosis	1	Major
Polydactyly	3	Major

Five of the listed malformations (marked with[a]) have been mentioned as typical for methimazole embryopathy in the literature (Clementi et al. 1999)

[a]Malformation regarded as typical for methimazole exposure

of cases in the cluster against the expected number. But the cluster should be presented so other investigators can check its validity in their material.

My conclusion is that one should always present the risk of any (at least major) malformation and then present the various malformations by a reasonable grouping, avoiding the "organ system" division. If the material is not too large, it is advisable to present a table with individual malformations (or combinations of malformations), followed with the number of such cases.

An example will be given on limited data after early pregnancy exposure to methimazole (Table 10.2). Only 151 infants were exposed, but 15 (10 %) of them had a congenital malformation diagnosis of variable significance. The risk ratio for any congenital malformation was 2.09 (95 % CI 1.17–3.45), for a relatively severe malformation 2.79 (95 % CI 1.52–4.68).

References

Bérard A, Iessa N, Chaabane S, Musanda FT, Boukhris T, Zhao J-P (2016) The risk of major cardiac malformations associated with paroxetine during the first trimester. Br J Clin Pharmacol 81(4):589–604. doi:10.1111/bcp.12849

Clementi M, Di Gianantonio E, Pelo E, Mammi I, Basile RT, Tenconi R (1999) Methimazole embryopathy: delineation of the phenotype. Am J Med Genet 83:43–46

Danielsson C, Brask J, Sköld A-C, Genead R, Andersson A, Andersson U et al (2013) Exploration of human, rat, and rabbit embryonic cardiomyocytes suggests K-channel block as a common teratogenic mechanism. Cardiovasc Res 97:23–32

Gardner JS, Guyard-Boileau B, Alderman BW, Fernbach SK, Greene C, Mangione EJ (1998) Maternal exposure to prescription and non-prescription pharmaceuticals or drugs of abuse and risk of craniosynostosis. Int J Epidemiol 27:64–67

Hernández-Dias S, Werler MM, Walker AM, Mitchell AA (2000) Folic acid antagonists during pregnancy and the risk of birth defects. N Engl J Med 343:1608–1614

Hernández-Diaz S, Werler MM, Walker AM, Mitchell AA (2001) Neural tube defects in relation to use of folic acid antagonists during pregnancy. Am J Epidemiol 153:961–968

Källén B, Borg N, Reis M (2013) The use of central nervous system active drugs during pregnancy. Pharmacoceuticals 6:1221–1286

Matok I, Gorodischer R, Koren G, Landau D, Wiznitzer A, Levy A (2009) Exposure to folic acid antagonists during the first trimester of pregnancy and the risk of major malformations. Brit J Clin Pharmacol 68:956–962

Olshan AF, Faustman EM (1989) Nitrosatable drug exposure during pregnancy and adverse pregnancy outcome. Int J Epidemiol 18:891–899

Timing of Drug Use and Effects on the Embryo or Fetus

11

The timing of the drug use is often crucial for a harmful effect on the embryo or fetus. This is perhaps most evident for the origin of congenital malformations. For these, the first trimester of pregnancy is of greatest importance. For other outcomes the period of sensitivity is less strict even though it is often thought that the strongest effects on, for instance, preterm birth and neonatal morbidity will be obtained by exposures during the second and third trimester.

The actual time of drug exposure is often uncertain, less so when it is based on interview data than when it is based on prescription data. In a few situations, drugs taken long before pregnancy may remain in the body and cause damage during pregnancy. An example of a drug which disappears very slowly from the patient is the teratogenic retinoid etretinate. Patients who had used that drug are recommended not to get pregnant for a period of 2 years after stopping the drug. This is unusual; in most instances drugs are rapidly metabolized or excreted and their half-life relatively short.

11.1 Exposure Before Conception

The use of a drug before conception could affect sperm or egg cell formation resulting in an abnormal development of the embryo after conception.

Direct damage of eggs or sperm should most likely cause gene mutations or chromosome anomalies. Candidates for such effects are mainly drugs with known mutagenic effects, notably drugs used for treatment of cancer or as immunosuppressant drugs. The typical result of a mutagenic effect would be a condition which is due to a dominant mutation, e.g., achondroplasia. The spontaneous rate of this condition is about 1/10,000 births, and a very strong increase would be needed in order to detect it in a material consisting even of a few thousand exposed pregnancies.

Note that in order to get a mutagenic effect which causes a malformation, the exposure most likely must occur before or around conception. Mutagenesis which takes place during the fetal development may increase the risk of childhood cancer but is unlikely to cause a malformation.

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_11

Sperm damage leading to infertility occurs with some chemicals, including some drugs, for instance, methotrexate and finasteride. Such effects are usually temporary and disappear a time after stopping the drug. Another possibility is that the drug causes a mutation in the sperm which will participate in the conception. A third possibility, suggested in animal experiments, is that the drug can cause epigenetic modifications (Cordier 2008).

The most critical period for mutations in sperms is about 3 months before conception, while mutations in eggs can occur any time since the woman's birth which makes studies of the phenomenon very difficult. So far studies of women who had been treated with potentially mutagenic drugs or radiation because of childhood malignancy and later had pregnancies have not demonstrated an increased risk for malformations in the offspring.

Paternal exposure to drugs like azathioprine/6-mercaptopurine has not been linked to an increased risk for congenital malformations (e.g., Hoeltzenbein et al. 2012, but the studied number of exposures was low, $n=115$). In a paper from Motherisk (Lee et al. 2010), 301 TIS questions on paternal exposures were tabulated. They referred both to drugs with a potential to mutagenesis (methotrexate, azathioprine) and drugs with a known or suspected teratogenic activity when used by the woman (e.g., isotretinoin, valproic acid). Among 43 live births, one infant had a congenital malformation, preaxial polydactyly.

11.2 Period of Organogenesis: First Trimester Exposure

Most congenital malformations are a result of a disturbed organogenesis, that is, the formation of the organs during embryonic development. Generally speaking this occurs during the first trimester, but a few congenital malformations can be formed later during pregnancy. Examples are microcephaly and hydrocephaly which may origin late in pregnancy and even postnatally. Some other malformations like pyloric stenosis are also thought to develop late. It is also possible that the growth of the organ (which occurs throughout the whole pregnancy) can be affected with hypoplasia as the result.

In some animal experiments, one has seen teratogenic effects to occur before embryo implantation. Already during the passage of the egg through the fallopian tube, an exchange of chemicals between the maternal organism and the embryo may occur. Such very early damage will probably result in embryonic death and will go unnoticed clinically.

The main pregnancy period when a malformation can arise is, however, the first trimester. Within that period specific time windows exist for the origin of specific malformations, that is, periods when the relevant structures develop. Such "sensitive periods" may be the most relevant ones for teratogenesis and were very typically demonstrated for thalidomide. Theoretically, a drug treatment before the formative period of an organ can cause a malformation because every structure comes from an earlier rudiment. Even though the organogenesis of the heart starts during week 5, the first cardiogenic areas exist already toward the end of the third week and damage at that time may result in later disturbances of heart morphogenesis.

Table 11.1 Approximate timetable of organogenesis for some common malformations

Malformation	Weeks
Anencephaly	3–4
Spina bifida	4
Encephalocele	3–4
Holoprosencephaly	4
An- or microphthalmia	4–6
External ear malformations	4–5
Major heart malformations	4–8
Ventricular septum defects	6–10
Atrium septum defects	5–6
Choanal atresia	5–7
Cleft lip/palate	5–7
Median cleft palate	8–12
Esophageal atresia	5–6
Anal atresia	5
Diaphragmatic hernia	4–7
Omphalocele	6–10
Gastroschisis	5–6?
Major kidney malformations	5–12
Hypospadias	7–14
Limb reductions	4–5
Polydactyly	5–6
Syndactyly	5–6

Modified after Czeizel (2008)
The stated weeks refer to time after conception, not time after last menstrual period
[?]The exact timing of gastroschisis is unclear as its mode of embryogenesis is debated

A crude tabulation of the weeks of interest for some congenital malformations is given in Table 11.1. The weeks given are only approximate but can give an idea of the time pattern. What can be seen is that many malformations may be formed already so early that the woman hardly suspects her pregnancy.

If the drug under study is used during the whole first trimester – a common situation at chronic use of the drug – the period of sensitivity for a specific malformation is relatively uninteresting, but if the drug has been used only for a short time like a week (e.g., an antibiotic), a risk estimate based on exposure during the whole first trimester will be strongly biased toward null, even if we concentrate the study to exposures during the second and third month after LMP as has been suggested (Czeizel 2008).

Let us take an example: use of erythromycin has been associated with an increased risk of cardiovascular defects (Källén et al. 2005). This finding was based on 31 infants with such malformations among 1844 infants exposed to erythromycin in the first trimester which gave an OR =1.84 (95% CI 1.29–2.62). The first trimester is 10 weeks long (counted from conception) and we suppose that the cardiac rudiment is sensitive during 5 weeks. If erythromycin as an average was used

for 1 week, crucial exposure might have occurred during a total of 35 days of the sensitive period. Thus the actual number of infants exposed during the sensitive period would be $1844 \times 35/70 = 922$, and the actual risk at exposure during the sensitive period could be twice as high as that estimated. These are hypothetical calculations but show in what direction exposure outside the sensitive period will affect risk estimates.

If a drug has been used after the formative period of a malformation, it is unlikely to cause it. Some years after the thalidomide tragedy, a rumor was spread in Sweden of an association between the use of meclizine and spina bifida (as far as I know never published outside local media). Meclizine is an antihistamine mainly used for NVP which seldom starts before week 6. Extensive information exists on the harmlessness of meclizine in early pregnancy (Källén and Mottet 2003) – but nearly all data represent treatments of NVP and that would be too late to cause spina bifida. Only if meclizine had been used for other purposes, e.g., moving sickness, before the woman knew about her pregnancy, could an effect on the spina bifida rate occur.

In efforts to use exact timings of exposures during the first trimester in order to increase the sensitivity of the study, one often has the added problem of uncertainty in dating. Most interview or questionnaire information asking which week the drug was used relies on the woman's statement. If the information is obtained in early pregnancy, an uncertainty about pregnancy week may exist to which may be added an uncertainty if the woman counts from LMP, from conception, or from the first missed period. If the interview is made after delivery when the woman knows about the presence or absence of a congenital malformation, exact dating will add to the uncertainty of exposure data due to the possibility of recall bias.

Studies based on prescription registers can identify the earliest possible exposure but not the actual exposure time. The possibility that the woman had access to and used the drug before the data of filling the prescription adds to the uncertainty of time of exposure.

The possibility that teratogenic drugs could be transferred from the man to the woman at intercourse during early pregnancy and reach and damage the embryo is regarded as unlikely. This was discussed, for example, for isotretinoin which is a drug with a strong teratogenicity and recommendations existed to avoid conception during male therapy with this drug. Estimates have shown, however, that the transferred amounts are so small that it is unlikely that it could harm the embryo (Millsop et al. 2013). No such effect was known of paternal use of thalidomide.

11.3 Exposure After the First Trimester

With few exceptions such exposures will be irrelevant for the origin of congenital malformations but may cause other adverse outcomes like preterm birth, low birth weight, neonatal morbidity, and long-term effects. It is, however, possible that also early exposures may affect outcomes around delivery. Placentation and placenta development may be affected which could, for instance, increase the risk of placental abruption but also of preterm birth. A further consideration is that women who

used a drug in early pregnancy may be more likely than other women to use it also later in pregnancy even if she got no new prescription, so early use could be a proxy for later use.

In a recent study of the effect of air pollution of term infant birth weight (Rich et al. 2015), it was suggested that an effect was found only during the eighth month of pregnancy.

References

Czeizel A (2008) Specified critical period of different congenital abnormalities: a new approach for human teratological studies. Congenit Anom (Kyoto) 48:103–109

Cordier S (2008) Evidence for at role of paternal exposures in developmental toxicity. Basic Clin Phamracol Toxicol 102:176–181

Hoeltzenbein M, Weber-Schoendorfer C, Borisch C, Allignol A, Meister R, Schaefer C (2012) Pregnancy outcome after paternal exposure to azathioprine/6-mercaptopurine. Reprod Toxicol 34:364–369

Källén B, Mottet I (2003) Delivery outcome after the use of meclizine in early pregnancy. Eur J Epidemiol 18:665–669

Källén BAJ, Otterblad Olausson P, Danielsson BR (2005) Is erythromycin therapy teratogenic in humans? Reprod Toxicol 20:209–214

Lee CYW, Jin C, Mata AM, Tanska AM, Einarsson A, Koren G (2010) A pilot study of paternal drug exposure: the Motherisk experience. Reprod Toxicol 29:353–360

Millsop JW, Heller MM, Eliason MJ, Murase JE (2013) Dermatological medication effects on male fertility. Dermatol Ther 26:337–346

Rich DQ, Liu K, Zhang J, Thurston SW, Stevens TP, Pan Y et al (2015) Differences in birth weight associated with the 2008 Beijing Olympics air pollution reduction: results from a natural experiment. Environ Health Perspect 123:880–887

Repeated Studies and Meta-analyses 12

As pointed out in the text above, the findings of a single study may well be a chance finding in spite of a formal statistical significance. This should not prevent publication but further information would be needed before a final decision is made about the risk with the drug exposure. Such repetitions can be made by other scientists or by a new study by the original authors, using a different material.

We can illustrate this with a concrete problem – the possible association between maternal use of the antihistamine loratadine and an increased risk of hypospadias in the male offspring. When the Swedish Medical Birth Register began recording maternal drug use (on July 1, 1994), this was initiated by a committee which also contained representatives from the drug industry. When data had been collected for a couple of years, it was suggested that one should test the system using a relatively new antihistamine, loratadine, as the test drug. This was done and the first analysis was made in June 1997. At that time there were 355 exposed infants and the total number of congenital malformations was the expected one, but there were two cases of hypospadias (0.7 expected). This could obviously be a chance finding, and there was nothing known in the mechanism of action of the drug which suggested an antiandrogenic effect. The monitoring continued as seen in Fig. 12.1. In May 1998 there were seven infants with hypospadias against 1.6 expected and a real concern was raised about a possible causal association. The next 3 years, however, only two further cases were seen which supported the thought that the original cluster was random. Then a new outburst of cases occurred so in November 2001, there were 15 cases against the expected number of 5.6 – since the first observed cluster in May 1988, a further eight cases had occurred against the expected number of four. Numbers were small and the malformation not very serious, but it was felt necessary to report the finding (Källén and Otterblad Olausson 2001) even though it was stressed in the article that the finding could be random.

Relatively quickly studies were published from other scientists. Some small studies were published (Diav-Citrin et al. 2003; Moretti et al. 2003) but they had no power to detect a doubling of the rate of hypospadias (a total of 210 and 161 exposed infants, respectively). A third study came from CDC in Atlanta and was based on

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_12

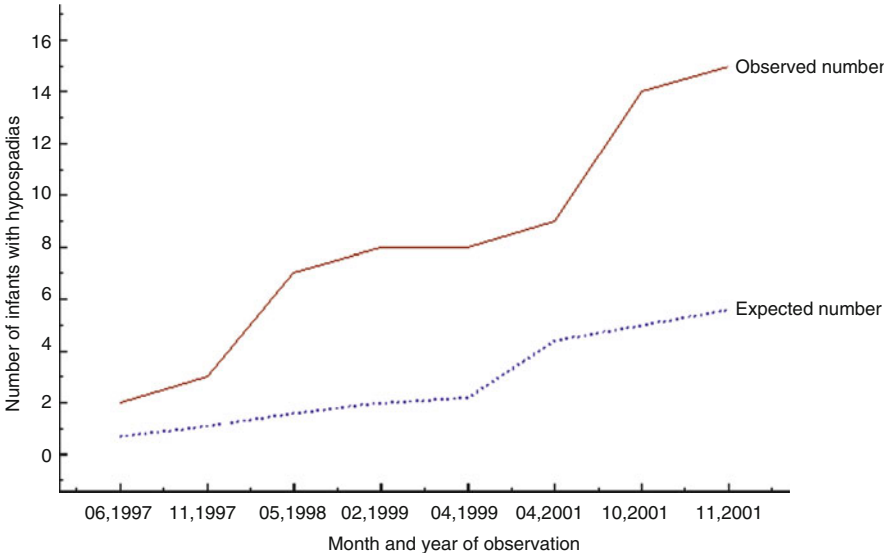


Fig. 12.1 Observed and expected numbers of hypospadias after maternal use of loratadine in early pregnancy at eight occasions (month, year) (Data from Källén and Otterblad Olausson (2001))

the National Birth Defects Prevention Study. In this study only penile or more severe hypospadias were studied, and in our material all but one case was of the most common type with the urethral opening in the sulcus coronaries of the penis. Two somewhat larger studies were published from Denmark (Pedersen et al. 2006a, b): one was based on linkage with a prescription register and the other on the prospective Danish National Birth Cohort. No increased risk for hypospadias was detected but the confidence intervals were wide and based on few exposures. Upper confidence limits were 10.4 and 6.9, respectively. These studies illustrate the difficulties to falsify a statement when it concerns a malformation – very large studies are needed.

In the meantime we had continued the monitoring of loratadine and hypospadias using further sources of malformation identification (Källén and Otterblad Olausson 2006). For the period 2002–2004, we identified 1911 infants exposed to loratadine – only two had hypospadias and the expected number was 4.6, and the RR was 0.47 with a 95 % CI of 0.06–1.68. The rates of hypospadias after loratadine exposure in the two periods were 25/2780 and 2/1911. These two rates differ significantly ($p<0.001$). Our conclusion was that most likely the high number during the first period of observation was due to the multiple testing situation which exists in the monitoring process.

This conclusion is supported by the continued monitoring. During the years 2005–2013, there were 4315 loratadine exposures and 11 cases of hypospadias, OR = 0.85 (95 % CI 0.47–1.53).

This example illustrates how even a strong association may arise by chance and how large materials are needed to eliminate a suspicion of causality. At the present

time, one can think it was unnecessary to publish the first cluster, notably as it referred to a malformation which was not very severe. A likely mechanism of action was unknown – no antiandrogenic effect was known of the drug – but one hypothesis was that the drug could affect the fetal testicles which could have had effects on the future reproductive capacity of exposed boys, something which would not be evident until decades later.

Another example refers to the observation of an association between maternal use of erythromycin and infant cardiovascular defects. This was first noticed in a study of maternal drug use and infant cardiovascular defects (Källén and Otterblad Olausson 2003) based on Swedish Medical Birth Register data from 1995 to 2001. The study tabulated 68 different drugs or drug groups and found 14 with a "statistically significant" increased risk and one with a decreased risk. Some of these were previously known or suspected associations like insulin, antihypertensives, fertility drugs, and anticonvulsants, but some were not known or suspected before, including macrolides and erythromycin. This association was scrutinized in a further paper (Källén et al. 2005) which contained data from one more year (2002). The OR for any cardiovascular defect was 1.84 (95 % CI 1.29–2.62) based on 31 exposed cases. Eighteen of them were ventricular or atrial septum defects. A hypothesis was presented to explain the association: a side effect of erythromycin is an inhibition of a specific cardiac potassium current (IKr) channel which according to animal experiments could result in a cardiac malformation. The possibility that the observation was a result of multiple testing was also stressed.

A follow-up study using the same source of data for a few years more (Källén K 2005) found a lower and nonsignificant OR, but the two estimates did not differ significantly. A few studies from other parts of the world could not verify the association (e.g., Bérard et al. 2015). Extended Swedish data were examined again in a paper by Källén and Danielsson (2014). The OR for a cardiovascular defect for the period 1996–2011 was 1.70 (95 % CI 1.26–2.29). When the observation period was divided into two halves (1996–2003, 2004–2011), the OR estimates were nearly identical: 1.69 and 1.71. What had happened in the meantime was that the use of erythromycin in early pregnancy had drastically decreased: from 2.7 per 1000 the first 8-year period to 0.7 per 1000 the second 8-year period. The RR for the second period was therefore not statistically significant (95 % CI 0.78–3.25). One thus had an estimate (1.71) which can be compared to two values. One is 1.0 which indicates no effect; the other is 1.69 which was the significant estimate for the first 8 years. Obviously, the second comparison is more relevant than the former. This way of reasoning is related to Bayesian statistics. It can be pointed out that one study who declared that erythromycin lacked teratogenic capacity actually registered an odds ratio of 1.6 which was not statistically significant from 1.0 but neither from our estimate of 1.7 (Romøren et al. 2012).

A more formalized method to analyze data from repeated studies of the same problem is a meta-analysis. The idea is simple: if a series of studies present risks which all are estimates of one true risk, one would get a better risk assessment if the various risks were pooled and weighted according to the size of the studies. In the ideal situation, the meta-analysis should be made on all studies (published or not)

which were performed with identical methodology; all studies should report the same experiment or type of observation. This is often the case in clinical studies which is also the situation where meta-analyses have played the largest role.

When epidemiological studies are to be compared, the basic prerequisites for a meta-analysis are seldom fulfilled. So, for instance, methods of ascertainment of drug use and of presence of malformations vary and various biases may be included differently in the studies. Typically one or a few studies are much larger than other studies and will dominate the common risk estimate – which means that the end result to a large extent will depend on the methodology and quality of these large studies.

As a part of the meta-analysis, a selection of studies based on quality is usually made, dismissing, for instance, studies without controls. In spite of formal guidelines, it will end with the subjective idea of the persons who select data for the analysis. Let us take a simple example: the US National Birth Defects Research Program is probably regarded by many as being of a high standard (otherwise results would not get published in major journals) in spite of the fact that exposure data are retrieved retrospectively and that the percentage of nonresponders is high, about 30 %, facts which make other researches regard their results with suspicion.

Unfortunately, the fact that a paper is published in a large medical journal does not mean that the results are believable; sometimes it only means that they are spectacular!

Personally, I think that a detailed discussion of the pros and cons of the individual studies is more important than to get a weighted common risk estimate. If large studies differ in results, this may more likely be an effect of design than of random variation around the true risk. In the next chapter, I will summarize some questions which can be put when one wants to scrutinize the validity of a published study. A critical discussion of the results of large and methodologically acceptable studies should be carried out, and a conclusion can be drawn, perhaps without an effort to pinpoint a specific risk level.

References

Bérard A, Sheehy O, Zhao J-P, Nordeng H (2015) Use of macrolides during pregnancy and the risk of birth defects: a population-based study. Pharmacoepidemiol Drug Saf 24(12):1241–1248. doi:10.1002/pds.3900

Diav-Citrin O, Schechtman S, Abaronovich A, Moerman L, Amon J, Waijnberg R et al (2003) Pregnancy outcome after exposure to loratadine or antihistamines; a prospective controlled cohort study. J Allergy Clin Immunol 111:1239–1243

Källén B, Danielsson B (2014) Fetal safety of erythromycin: an update of Swedish data. Eur J Clin Pharmacol 70:355–360

Källén B, Otterblad OP (2001) Monitoring of maternal drug use and infant congenital malformations. Does loratadine cause hypospadias? Int J Risk Safety Med 14:115–119

Källén B, Otterblad Olausson P (2003) Maternal drug use in early pregnancy and infant cardiovascular defect. Reprod Toxicol 17:255–261

Källén B, Otterblad Olausson P (2006) No increased risk of infant hypospadias after maternal use of loratadine in early pregnancy. Int J Med 3:106–107

Källén B, Otterblad Olausson P, Danielsson B (2005) Is erythromycin therapy teratogenic in humans? Reprod Toxicol 20:209–214

Källén KBM (2005) Maternal drug use in early pregnancy and risk of cardiovascular defect in infant – a follow-up study. Reprod Toxicol 26:64–65

Moretti ME, Caprara D, Coutinho CJ, Bar-Oz B, Berkovitch M, Addis A et al (2003) Fetal safety of loratadine use in the first trimester of pregnancy: a multicenter study. J Allergy Clin Immunol 111:479–483

Pedersen L, Vinther Skriver M, Nørgaard M, Sørensen HT (2006a) Maternal use of loratadine during pregnancy and risk of hypospadias in offspring. Int J Med Sci 3:21–25

Pedersen L, Nørgaard M, Vinther Skriver M, Olsen J, Sørensen HT (2006b) Prenatal exposure to loratadine in children with hypospadias: a nested case-control study within the Danish National Birth Cohort. Am J Ther 13:320–324

Romøren M, Lindbæk M, Nordeng H (2012) Pregnancy outcome after gestational exposure to erythromycin – a population-based register study from Norway. Br J Clin Pharmacol 74:1053–1062

The Identification of Risks and the Information Problem

13

13.1 Pharmacovigilance

The main type of pharmacovigilance – to monitor adverse events of drug use – is based on voluntary reporting of such events. This activity was discussed in Chap. 3 and has been valuable in some situations, notably when new drugs were introduced on the market. In case of a very strong teratogenicity (for instance of equal strength as that of thalidomide), already a few reports of severe malformations could result in the identification of the hazard, especially if the reported malformations showed similarities in their pathogenesis.

There is, however, a need for other ongoing monitoring of the effects of maternal use of drugs on pregnancy outcome, not relying on voluntary reporting. In 2003 Mitchell pointed out that the American FDA should develop a teratogen surveillance system. The Swedish Medical Birth Register gives opportunities for such monitoring. In contrast to most scientific studies which are limited in time, this register allows an ongoing monitoring of the outcome after drug use during pregnancy. Up till now this has been performed in the following way which certainly is not ideal but which perhaps can give the reader some ideas.

Data from the register concerning drug use are updated once a year. At this updating, the drug names as written down by the midwives who performed the interviews are transferred to ATC codes. The complete register is then analyzed. All ATC codes are tabulated and drugs or drug groups with at least 50 exposures are identified. For each drug (group) the number of congenital malformations as recorded in the Medical Birth Register is calculated, and the expected number is compared with the observed number with a Poisson distribution model or with chi-square tests if numbers are large enough. "Significant" differences (as increases or decreases) are identified. The process is repeated for a number of malformation groups: all malformations, relatively severe malformations, neural tube defects, cardiovascular defects, orofacial clefts, alimentary tract atresia, hypospadias, etc. No adjustment for possible confounders is made but only crude numbers and risks are

B. Källén, *Drugs During Pregnancy*, DOI 10.1007/978-3-319-40697-8_13

given. A special interest is given to risk estimate changes from the previous years, appearance or disappearance of risks.

At a meeting with representatives from the Swedish National Board of Health and Welfare and the Medical Drug Agency, the list is scrutinized and discussed against the background of what is known from the literature. When something noteworthy is seen, a detailed investigation is made and more complete information on malformations is collected, and the possible effect of confounding is considered.

This system would miss a strong teratogenicity of a new drug, perhaps used by less than 50 women but with a high risk of congenital malformations (the thalidomide type). Furthermore, no data on other pregnancy outcomes or long-time effects are analyzed.

Specific investigations may also be initiated from literature reports or from adverse drug reaction reports. This type of monitoring should be a natural consequence of running a register. The authority which runs the register should not only allow scientists access to the data but should use the data for an ongoing monitoring of drug use during pregnancy.

Another system for "signal detection" was described from the Netherlands (de Jonge et al. 2013) where data from a malformation register with respect to drug exposure (based on pharmacy records, verified by telephone interviews after delivery) were compared with data from a population register (where drug exposure was identified from a prescription register). It is not clear if this has developed into an ongoing surveillance or if it only was a way to test a possible approach.

13.2 Information on the Risk with Drug Use During Pregnancy

In an ideal world, every published scientific article should give a clear and unequivocal answer to the question posed. As has repeatedly been pointed out in the text above, this is seldom the case in the real world. Each research result should be looked upon as a piece in a jigsaw puzzle – the complete answer is not obtained until many pieces have been added, and a picture begins to appear which at least resembles the truth. The uncritical spreading of findings through various media and the Internet can be harmful and cause unnecessary anxiety. On the other hand, an event similar to the thalidomide tragedy would need a rapid spread of information to prevent further damage. A definite problem is that if the media regularly cry "Wolf!", finally a serious warning may not be believed by the public, as was the case in the legend of Aisopos.

One often sees a general recommendation – not to use drugs during pregnancy. This is a rather impractical rule. It is true that drugs which are not needed should not be used during pregnancy (or otherwise) and that, for instance, the use of street drugs by the pregnant woman should be strongly discouraged. The pregnant woman, however, must get adequate therapy when that is needed; the problem is to balance the need of the woman against the possible risk for the embryo or fetus. Lack of treatment may actually be a larger risk for the embryo than adequate drug use, for instance, at maternal asthma or epilepsy.

In the daily clinical work, questions on drug use during pregnancy are common and have to be answered. It is not reasonable to expect that practitioners and midwives should be able and have the time to follow and interpret the scientific literature on the subject. They would have to rely on digested impartial information. There have been many efforts to supply the medical practitioner with evaluated information. The first effort was probably made in Sweden where in 1978 a classification system was introduced in the Swedish National Drug Formulary (FASS) (Berglund et al. 1984). The classification of a drug is made by the drug industry. The system consists of the following groupings (somewhat abbreviated):

A. Drugs which have been much used and as far as known have no ill effects on the pregnancy or embryo. No consideration is taken of possible effects in animal tests.
B. Drugs which have been less used and where not enough experience of human use exists. This group is divided into three subgroups:
   B:1 Reproduction toxicology studies on animals have given no reason for concern.
   B:2 Reproduction toxicology studies on animals are incomplete but the available data give no reason for concern.
   B:3 Reproduction toxicology studies on animals have shown effects on the reproduction but their relevance for the human is unclear.
C. Drugs which may be harmful for the fetus or the newborn but without causing congenital malformations.
D. Drugs which in the human can cause malformations or are likely to cause malformations.

Soon after, a similar classification was introduced in the USA (FDA Pregnancy Categories) as follows (also somewhat abbreviated):

A. Adequate and well-controlled studies have failed to demonstrate a risk for the fetus.
B. Animal reproduction studies have failed to demonstrate a risk to the fetus, and there are no adequate and well-controlled studies in pregnant women.
C. Animal reproduction studies have shown an adverse effect on the fetus, and there are no adequate and well-controlled studies in pregnant women. Potential benefits may warrant use of the drug in pregnant women in spite of the potential risk.
D. There is positive evidence of human fetal risk, but potential benefits may warrant use of the drug in pregnant women in spite of the potential risk.
X. Studies in animals or humans have demonstrated fetal abnormalities, and the risk involved in the use of the drug in pregnant women clearly outweighs potential benefits.
N. Drug not classified.

Other such classifications have been made, for instance, in Australia.

The problem with these and similar classifications is that they are rather blunt and will sometimes be of little help in the practical situation. There are two major problems:

1. The woman may already have used the drug during pregnancy, and the question now is: should she interrupt her pregnancy – if not, can she continue to use the drug? For a determination of a pregnancy interruption because of a drug use, strong evidence for a substantial and serious risk must exist which is rare – an example is isotretinoin and of course thalidomide. In some situations the drug use may be a reason for intensified prenatal diagnosis.
2. Drugs are sometimes used because of a serious disease which may threaten the health of the woman, and this may in itself carry a risk for the fetus. An example is severe depression where the necessity to treat the woman may mean that the risk for an increased neonatal morbidity has to be accepted. Another such situation is a newly detected cancer in a pregnant woman needing cytostatic therapy – if this can be postponed till after the first trimester, the risk for the fetus is reduced.

For the evaluation of the situation for the individual patient, more specific information than a general pregnancy labeling is needed. The evaluation of a clinical pharmacologist may deviate from the crude classification (Erdeljic et al. 2010). Such information can be given by a TIS organization when such is available. Other efforts have been made, e.g., the Internet-based Janusinfo (in Swedish) which summarizes data from the Swedish Medical Birth Register and the scientific literature on individual drugs (Nörby et al. 2013). The site is written for medical professionals but has been much used also by the general public. It refers to the legal and health system in Sweden and may not be directly applicable in other countries.

An important issue is at what time during pregnancy the question turns up. A consultation of a woman who is planning a pregnancy and wonders if she can use a specific drug is reasonably simple. One should try to change drug therapy if a drug with definite or suspected hazards for the pregnancy and infant is used, e.g., valproic acid. When less strong and perhaps uncertain effects exist, like SSRI drugs, it may be possible to try to stop the drug use or at least to select the drug with the best track record.

In the most common situation, the woman turns up for consultation when she is in early pregnancy and then tells that she has already used a specific drug. Is this harmful for the embryo? If the question concerns a drug which has been studied in large and well-performed studies and has appeared harmless, obviously clear information should be given that the use of it does not endanger the baby. In this situation there is a not uncommon complication – the patient may search for information on the Internet and find a study which indicates a risk. One then has to explain why one does not believe in the published results. Perhaps this book can help.

If, on the other hand, a risk has (more or less certainly) been demonstrated, its significance has to be evaluated. If an exposure has already occurred, the question

should be reformulated to the following: Is the risk so large and the damage so serious that an interruption of the pregnancy should be considered? This is a relatively rare situation, but on the other hand, it is likely that many pregnancies are unnecessarily interrupted because of anxiety for low or nonexisting risks. In some instances, a detailed search for congenital malformations by fetal diagnosis may be recommended because of a drug use. A typical example is use of valproic acid which – among other effects – increases the risk for spina bifida some 10–20 times, and this malformation can often be detected prenatally and a pregnancy interruption can be performed if the woman so wishes. Other ill effects of this drug can often not be detected prenatally but may be less severe.

One sometimes hears the woman saying that she does not want to take any risk that the baby will be malformed, but she should know that the risk always exists and if this risk is marginally increased, it is of little practical importance.

In most instances, the question of a continued use of the drug during pregnancy does not refer to the risk for malformations but to other fetal risks. We can illustrate this with two examples.

A woman with essential hypertension has become pregnant when using an ACE inhibitor. The continued use of the drug during pregnancy represents a risk for intrauterine death of the fetus, and the drug should be stopped and other antihypertensive drugs should be used.

A woman with depression has become pregnant using an SSRI drug. Continued use of this drug during pregnancy will increase the risk for neonatal morbidity which as a rule is temporary. If she stops using the drug, the risk for a worsening of her depression is considerable and includes an increased risk for post-delivery depression. It may be better to accept the infant risk. If, however, her mental condition is such that a trapping down of the treatment can be made, this is to be preferred, among other things because of a possible effect on the long-term development of the child.

As pointed out above, studies of the effect of drugs on the long-term development of the child are difficult to perform in an adequate way. Such effects are important and are most likely to occur after the use of psychoactive drugs. If such exposures can be avoided, this is obviously best, but it is a difficult balance between the (often hypothetical) risk of the child and the sometimes strong need of the woman of drug therapy. Even though maternal use of opioids during late pregnancy increases the risk of neonatal abstinence syndrome (NAS), it is possible that long-term effects seen in children of opioid-abusing women less are due to intrauterine drug exposure than to postnatal environmental factors (Baldacchino et al. 2014; Sithisarn et al. 2012). The well-known consequences of high alcohol exposure during pregnancy, resulting in fetal alcohol syndrome which includes also long-term neuropsychiatric effects, is a warning that such risks may exist also with drug exposure. To balance these risks against the needs of the woman is very difficult and necessitates careful considerations and access to data from well-performed studies. There is also a need of competent interpretation of published data. I will finish with some suggestions for such an evaluation.

13.3 Concluding Remarks: How to Evaluate a Published Study

For the evaluation of studies related to maternal drug use and pregnancy or neonatal complications, some important things have to be considered.

Is the study large enough to be able to detect a moderately increased risk for a serious outcome? To study malformation risks in a material of one or two hundred exposed infants can only reveal very strong risk increases, however carefully the study is made. Such small studies are of limited value and cannot prove a lack of teratogenic properties in the drug.

Does the study have the power to investigate effects on specific malformations? Even a study comprising thousands of exposed infants may not be large enough to study effects on rare malformations like spina bifida or gastroschisis.

Are the data obtained retrospectively or prospectively related to the outcome studied? Retrospective studies where drug exposure is ascertained after the appearance of the harm (e.g., birth of a malformed child) are apt to falsely identify or exaggerate risks. Much consideration should not be paid to this type of study.

Is drug exposure ascertained from prescription registers? Exposure data from prescription registers may bias the risk estimates toward null and may therefore be unable to identify moderate risks, but they can hardly give falsely increased risks.

Have adequate considerations been taken of possible confounding? The fact that adjustment has been made for a number of different variables does not necessarily mean that significant confounders have been adjusted for. Are the variables which have been adjusted for really confounders and not mediators? Is underlying disease a likely cause of the outcome studied?

How many statistical tests have been performed? Is a finding the result of a "fishing party" or is it a test of a hypothesis which was set up in advance of the data collection?

In Studies of Congenital Malformations

Is the malformation rate in a control group reasonable? If the rate of major malformation is below 2 % or above 5 %, ascertainment is probably inadequate. A very low rate indicates a poor ascertainment, while a very high rate indicates inclusion of minor or uncertain anomalies which may hide effects on truly major malformations.

Are malformations described in adequate detail so possible effects on specific conditions can be identified? Studies which report only unspecified congenital malformations or birth defects are of less value than studies which describe specific malformation types. Are the malformation descriptions based on information from specialists (pediatricians, child pathologists, geneticists, etc.) or from general practitioners or parents?

References

Baldacchino A, Arbuckle K, Petrie DJ, McCowan C (2014) Neurobehavioral consequences of chronic intrauterine opioid exposure in infants and preschool children: a systematic review and meta-analysis. BMC Pscyhiatry 14:104. http://www.biomedcentral.com/1471-244X/14/104

Berglund F, Flodh H, Lundberg P, Prame B, Sannerstedt R (1984) Drugs used during pregnancy and breast-feeding. A classification system for drug information. Acta Obst Gynecol Scand Suppl 126:1–55

De Jonge L, Zetstra-van der Woude PA, Bos HJ, de Jong-van den Berg LTW, Bakker MK (2013) Identifying associations between maternal medication use and birth defects using a case-population approach. An exploratory study on signal detection. Drug Saf 36:1069–1078

Edeljic V, Francetic I, Makar-Ausperger K, Liklic R, Radafic-Aumiler M (2010) Clinical pharmacology consultation: a better answer to safety issues of drug therapy during pregnancy? Eur J Clin Pharmacol 66:1037–1046

FDA Pregnancy Categories. www.drugs.com/pregnancy-categories.html

Janus Info. www.janusinfo.se/beslutsstöd/lakemedel-och-fosterpaverkan/

Mitchell AA (2003) Systematic identification of drugs that cause birth defects – a new opportunity. N Engl J Med 349:2556–2559

Nörby U, Källén K, Eiermann B, Korkmaz S, Winbladh B, Gustafsson LL (2013) Drugs and birth defects: a knowledge database providing risk assessments based on national health registers. Eur J Clin Pharmacol 69:889–899

Sithisarn T, Granger DT, Bada HS (2012) Consequences of prenatal substance use. Int J Adolesc Med Health 24:105–112