British Journal of Clinical Pharmacology

Br J Clin Pharmacol (2016) **82** 1110–1122 1110

PHARMACOEPIDEMIOLOGY

EUROmediCAT signal detection: a systematic method for identifying potential teratogenic medication

Correspondence Joan K. Morris, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ. Tel.: +44 (0) 20 7882 6274; Fax: +44 (0) 20 7882 6270; E-mail: j.k.morris@qmul.ac.uk

Received 23 January 2016; **Revised** 19 May 2016; **Accepted** 18 June 2016

Johannes M. Luteijn[1], Joan K. Morris[1], Ester Garne[2], Joanne Given[3], Lolkje de Jong-van den Berg[4], Marie-Claude Addor[5], Marian Bakker[6], Ingeborg Barisic[7], Miriam Gatt[8], Kari Klungsoyr[9], Anna Latos-Bielenska[10], Nathalie Lelong[11], Vera Nelen[12], Amanda Neville[13], Mary O'Mahony[14], Anna Pierini[15], David Tucker[16], Hermien de Walle[17], Awi Wiesel[18], Maria Loane[19] and Helen Dolk[20]

[1]Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK, [2]Paediatric Department, Hospital Lillebaelt, Kolding, Denmark, [3]Centre for Maternal, Fetal and Infant Research, Institute of Nursing and Health Research, Ulster University, Newtownabbey, UK, [4]Department of Pharmacy, Unit of PharmacoEpidemiology and PharmacoEconomics, Groningen, The Netherlands, [5]Division of Medical Genetics, CHUV, Lausanne, Switzerland, [6]University Medical Centre of Groningen, Groningen, The Netherlands, [7]Department of Medical Genetics and Reproductive Health, Children's Hospital Zagreb, Medical School University of Zagreb, Zagreb, Croatia, [8]Department of Health Information and Research, Guardamangia, Malta, [9]Medical Birth Registry of Norway, The Norwegian Institute of Public Health and Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway, [10]Department of Medical Genetics, Poznan University of Medical Sciences, Poznan, Poland, [11]Center for biostatistics and epidemiology, INSERM U1153, Paris, France, [12]Provinciaal Instituut voor Hygiene (PIH), Antwerp, Belgium, [13]IMER Registry (Emila Romagna Registry of Birth Defects), Center for Clinical and Epidemiological Research University of Ferrara and Azienda Ospedaliero- Universitaria di Ferrara, Ferrara, Italy, [14]Public Health Medicine, Health Service Executive, Cork, Ireland, [15]National Research Council (IFC-CNR), Institute of Clinical Pharmacology, Pisa, Italy, [16]Congenital Anomaly Register and Information Service for Wales, Public Health Wales, Swansea, UK, [17]Department of Genetics, EUROCAT Northern Netherlands, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, [18]Mainz Model Birth Registry, University Children's Hospital Mainz, Germany, [19]Centre for Maternal, Fetal and Infant Research, Institute of Nursing and Health Research, Ulster University, Newtownabbey, and [20]Centre for Maternal, Fetal and Infant Research, Institute of Nursing and Health Research, Ulster University, Newtownabbey, UK

Keywords adverse drug reactions, congenital anomalies, drug safety, pharmacoepidemiology, pharmacovigilance, pregnancy

AIMS

Information about medication safety in pregnancy is inadequate. We aimed to develop a signal detection methodology to routinely identify unusual associations between medications and congenital anomalies using data collected by 15 European congenital anomaly registries.

METHODS

EUROmediCAT database data for 14 950 malformed foetuses/babies with first trimester medication exposures in 1995–2011 were analyzed. The odds of a specific medication exposure (coded according to chemical substance or subgroup) for a specific anomaly were compared with the odds of that exposure for all other anomalies for 40 385 medication anomaly combinations in the data. Simes multiple testing procedure with a 50% false discovery rate (FDR) identified associations least likely to be due to chance and those associations with more than two cases with the exposure and the anomaly were selected for further investigation. The methodology was evaluated by considering the detection of well-known teratogens.

DOI:10.1111/bcp.13056

© 2016 The British Pharmacological Society

RESULTS

The most common exposures were genitourinary system medications and sex hormones (35.2%), nervous system medications (28.0%) and anti-infectives for systemic use (25.7%). Fifty-two specific medication anomaly associations were identified. After discarding 10 overlapping and three protective associations, 39 associations were selected for further investigation. These associations included 16 which concerned well established teratogens, valproic acid (2) and maternal diabetes represented by use of insulin (14).

CONCLUSIONS

Medication exposure data in the EUROmediCAT central database can be analyzed systematically to determine a manageable set of associations for validation and then testing in independent datasets. Detection of teratogens depends on frequency of exposure, level of risk and teratogenic specificity.

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Use of prescription medication in pregnancy is common.
- Since animal studies cannot accurately predict human teratogenesis and pregnant women are not involved in pre-marketing safety studies, little is known of the safety of medications in pregnancy
- Reproductive pharmacovigilance and, in particular, signal detection, is deficient for congenital anomalies.

WHAT THIS STUDY ADDS

- A novel methodology for systematic signal detection is able to identify potential medication congenital anomaly associations for validation and further investigation.
- Validated associations would require subsequent testing in hypothesis driven studies.

Introduction

Organogenesis occurs during the first trimester of pregnancy. This period is uniquely vulnerable to disruptive effects caused by teratogens, which can lead to congenital anomalies (CAs). Several prescription medications have been identified as being teratogenic, such as thalidomide [1] and valproic acid [2]. However, the aetiologies of most non-syndromic structural CAs remain incompletely understood and most are thought to result from a complex interplay between genetic, epigenetic, environmental and lifestyle factors [3]. Little is known about the safety of new medications in pregnancy. This is because, due to ethical considerations, pregnant women are not involved in pre-marketing medication safety studies. Additionally, medication safety studies in animals cannot accurately predict human teratogenesis. For example rodents are insensitive to isotretinoin, which is highly teratogenic in primates and humans [4]. Since 1980 the mean time taken to determine teratogenic risk for prescription medications approved by the Food and Drug Administration (FDA) has been 27 years [5].

Two studies estimated that 57% and 79% of pregnant women in the Netherlands and Norway, respectively, were exposed to prescription medications some time during pregnancy and 44% and 33%, respectively, in the first trimester [6, 7]. In addition to medication required for acute conditions, pregnant women with chronic diseases such as epilepsy, asthma and severe depression need to take their medication. Women may also inadvertently take medication during early pregnancy as an estimated 50% to 60% of European pregnancies are unplanned [8].

Reports on many teratogenic medications have been from spontaneous reports on specific CAs, rather than systematic pro-active surveillance [9]. For example thalidomide was brought to the attention of the scientific community by Dr McBride in the Lancet [10]. These suspected teratogens have then been investigated with case–control study designs.

To attempt to identify teratogens in a timely manner, international hypothesis generating studies on spontaneous reports have been performed using the Uppsala Monitoring Centre's VigiBaseTM [11], the European Medicines Agency's EudraVigilance [12] and the FDA's Adverse Event Reporting System (FAERS) [13] databases. All possible medication–adverse medication reaction combinations are considered and data mining algorithms have been developed using the proportional reporting ratio (PRR), the reporting odds ratio (ROR), the information component (IC) and the empirical Bayes geometric mean (EBGM) to identify those most likely to be true associations (signals) [13].

Spontaneous reports are limited by a number of factors including potential duplicates, large scale underreporting and biased reporting (e.g. after media attention) [14–16]. The underreporting of adverse medication reactions to spontaneous reporting systems has been estimated at 94% (95% CI 82, 98%) with no differences detected between general practice and hospital based studies [17]. These systems also often lack detailed information on CAs or the gestational age at which the medication was taken.

An alternative method of identifying teratogens is to use patient registries, for example EURAP (an international registry of anti-epileptic medications and pregnancy in which pregnancies to mothers taking anti-epileptic medications are registered before 16 weeks gestation and their outcomes followed-up) [18], to compare the risk of major CAs following maternal intake of different medications. These registries do not have the same issues of bias as the spontaneous report

systems as women are registered before the outcome of pregnancy is known, but their results are limited to certain medications or classes of medications.

CAs are rare, affecting approximately 2–3% of new-borns. The prevalence of specific CAs is even rarer with even well-known anomalies such as spina bifida affecting less than five per 10 000 pregnancies [19]. Large databases that cover millions of births are therefore necessary for meaningful analysis exploring the aetiology of CAs.

The aim of EUROmediCAT is to build a European population-based reproductive pharmacovigilance system. Such a system needs to be capable of signal detection and signal strengthening and evaluation, in a timely manner. We report here the development of a systematic signal detection algorithm to identify potential signals. This builds on previous research [20], but we apply here a statistical methodology to deal with the problem of multiple testing, analyze a large and high quality database (including also terminations of pregnancy for foetal anomaly) and look systematically across all medication exposures recorded rather than common exposures or exposures of prior interest. We report here the statistical methodology and assess its validity by comparing the results with known associations already reported in the literature [21]. The validation of the individual signals found and their correspondence with current scientific knowledge is reported in an accompanying paper [22].

Methods

Study population and database

EUROmediCAT is built upon the European Surveillance of Congenital Anomalies (EUROCAT) network [23]. EUROCAT registries are population-based and record all CAs occurring in foetuses that result in a live birth, foetal death from 20 weeks of gestation or termination of pregnancy following prenatal diagnosis. All the members of EUROCAT use similar inclusion criteria and have a consistent approach to data collection, coding and recording, which is monitored using data quality indicators [24]. Up to nine CAs coded according to the International Classification of Diseases (versions 9 and 10) and an unlimited number of medications taken during the first trimester coded up to seven digits (97%) using the Anatomical Therapeutic Chemical (ATC) classification can be registered. The ATC classification system is a WHO-controlled hierarchical medication classification system. ATC codes can be up to 7 digits, with the first five digits (ATC-4 level) representing chemical subgroup and all seven digits (ATC-5 level) representing chemical substances. Text information is available for each anomaly and each medication exposure. Additional data collected for CA registrations is described in detail in EUROCAT Guide 1.3 [25].

The signal detection dataset included all foetuses with CAs (excluding genetic conditions [25]) exposed to at least one medication *in utero* (excluding folic acid, minerals and/or vitamins) born between 1995 and 2011 from 15 registries in 13 European countries (Table 1), which participated in EUROmediCAT. Foetuses with isolated congenital dislocation of the hip were also excluded since the aetiology is mechanical, rather than potentially teratogenic.

Foetuses with only the following medication exposures were excluded from the study:

1. Medications not coded up to ATC-4 level (i.e. with <five digits)
2. Topical medications: S01-S03, D01A, D02-D04, D05A, D06-D09, D10A, D11AA, D11AC, D11AE, D11AF, D11AH01-D11AH03, M02 and all D11AX codes except for oral preparations.
3. Medications specified as occurring in the second/third trimester or with unknown timing.Ethical approval for this study was provided by the Ulster University Nursing Research Governance Filter Committee.

Cases and controls

For each specific CA, cases were foetuses with the specified CA from one of 57 EUROCAT pre-defined CA subgroups [25] and controls malformed foetuses without the specified CA. In addition to the 57 subgroups the EUROCAT coding committee recommended analysis of two new groups: laterality defects (atrial isomerism, dextrocardia, situs inversus, bronchopulmonary isomerism and asplenia and polysplenia) and neural crest defects (coloboma, Hirschprung's disease, Pierre Robin sequence, Goldenhar's syndrome, double outlet left/right ventricle, truncus arteriosus, interrupted aorta, double aortic arch, transposition of great arteries, Fallot's tetralogy and Arnold Chiari I/II malformations).

Exposure data

Registries collect medication exposures occurring during the first trimester of pregnancy, defined from the first day of last menstrual period up to the twelfth week of gestation [25]. Information on medication exposure was obtained mainly from obstetric/midwife records created before birth [26]. Additional data sources available for some registries included the medical records of the infant, records from the general practitioner, maternity passports and maternal interviews before or after birth [26]. In the northern Netherlands prescription data were used as an additional data source [26]. For Norway, the data normally transmitted to EUROCAT [26] were replaced by first trimester exposure from prescription redemption records. The EUROmediCAT project coded all medications from 1995 to ATC codes (previously this had only been done from 2005). Analysis was performed on ATC-4 exposures (chemical subgroup) and ATC-5 exposures (chemical substance group). ATC-4 codes were only considered if they yielded more exposures than corresponding ATC-5 codes. Information on dose and duration of medication exposure was not available for all registries. Medication exposure would include the woman taking a medication only once during the first trimester.

Substances that can be coded using multiple ATC-codes were identified using a WHO supplied list and the multiple ATC-5 codes were replaced with the single substance name. ATC codes subject to alterations over time were retrieved from the WHO website [27]. For all ATC-5 alterations, the old and new ATC-codes were replaced with the single substance name. No ATC-4 codes in the dataset were subject to ATC alterations. ATC alterations with special notes were not considered and in the event that a single ATC-code was

Table 1

EUROmediCAT signal detection dataset

EUROCAT registry	Birth years enrolled	Exposed foetuses with congenital anomalies (*n*)	Foetuses with congenital anomalies following data cleaning by timing of exposure[a] (*n*)	Data loss by data cleaning (%)	Total eligible ATC-coded exposures (*n*)	Average number of ATC-coded medication exposures per pregnancy
Belgium, Antwerp	1997–2011	358	354	1	529	1.49
Croatia, Zagreb	1995–2010	184	180	2	228	1.27
Denmark, Odense	1995–2011	234	234	0	357	1.53
France, Paris	2001–2011	659	659	0	968	1.47
Germany, Mainz	2005–2011	142	139	2	158	1.14
Ireland, Cork and Kerry	1996–2009	259	258	0	355	1.38
Italy, Emilia Romagna[b,c]	1995–2011	2322	2322	0	3826	1.65
Italy, Tuscany	1995–2011	1082	1043	4	1418	1.36
Malta	1996–2011	298	297	0	445	1.50
Netherlands, North Netherlands	1995–2011	2374	1844	22	3036	1.65
Norway	2005–2010	3052	3052	0[d]	5537	1.81
Poland (excluding Wielkopolska)	1999–2010	11 997	1958	84	2450	1.25
Poland, Wielkopolska	1999–2010	2713	409	85	552	1.35
Switzerland, Vaud	1997–2011	298	294	1	435	1.48
UK, Wales	1998–2011	1907	1907	0	2807	1.47
Total	1995–2011	27 879	14 950	46	23 101	1.55

[a]Exclusion of CA registrations with only medication exposures of unknown timing
[b]During the period 1995 to 2004 Emilia Romagna database had space for only 5 medications to be recorded
[c]Terminations of pregnancy for foetal anomalies were excluded from the Emilia Romagna registry as information on medications is only available for live and still births
[d]For Norway, the data normally transmitted to EUROCAT were replaced by CAs linked to first trimester prescription redemption records only

linked to multiple substances, the ATC-code was not replaced with either substance name. Medications with three or more exposed foetuses were investigated. A total of 693 unique ATC-4 and ATC-5 exposures satisfied the criteria described above and reached the threshold of three or more exposed foetuses. Foetuses only exposed to medications that did not reach this threshold are still included as controls.

Statistical analysis

Statistical analysis – Fisher's exact test. Analysis was performed by comparing the odds of exposure of the cases to the specific medication to the odds of exposure of the controls to the same medication using a one-sided Fisher's exact test. One-sided tests were considered appropriate as the study population includes only foetuses with a CA and if a medication reduces the risk of a CA, by definition there will be fewer cases in the study population and hence a low power to detect preventive medications. Any preventive associations identified were not examined as the study purpose was screening for teratogens. For each of the 40 887 analyses (59 unique congenital anomalies multiplied by 693 unique medications), registries without exposures and registries without cases were excluded for that analysis. A simple, non-adjusted Fisher's exact test was utilized to avoid overparametrized models (37 804 out of 40 887 combinations involved below three exposed cases for that specific anomaly) and since its simplicity suits the purpose of screening for, as opposed to supporting, teratogenicity. It was decided that adjusting for registry would exclude some analysis with small numbers of exposed cases and therefore the registry effect would be considered once signals had been identified. Outcomes are reported in *P* values and odds ratios (OR). For analysis of ATC-5 exposures, all pregnancies exposed to only the ATC-4 code corresponding to the first five digits of the ATC-5 code under analysis were excluded from cases and controls as the specificity of exposure could not be determined.

Statistical analysis – multiple testing. A total of 59 CA subgroups and 693 medication groups were considered for 40 887 potential analyses. There were no data to perform 502 analyses after excluding registries without a specific exposure and without a specific anomaly. Therefore a total of 40 385 analyses were performed.

The usual tests at a 5% level of statistical significance mean that 5% of all medication–CA combinations will have a statistically significant result by chance alone. For instance if 10 000 tests are performed an expected 500 will be statistically significant even if there are no true associations. If there is, for example, one true association amongst all these tests then it will be difficult to determine which is the true association amongst the 501 positive results. To overcome this problem, the false discovery rate (FDR) was controlled by applying the Simes multiple testing procedure [28]. The FDR is the proportion of false positives among total positives. In the above example, the FDR would be over 99% (500 out of 501). A pilot study based on a 25% sample of cases exposed to medications determined that a FDR of 50% was appropriate. Statistical analysis was performed using Stata 12.1 [29].

Medication–CA combinations with less than three exposed cases were included in multiple testing, but not considered for follow-up as signals. These associations were retained in the multiple testing procedures as lack of low powered associations violates the underlying assumptions of the multiple testing procedures by strongly shifting the distribution of p-values towards zero.

Follow-up of signals. Signals belonging to medication groups which were being separately investigated as part of the EUROmediCAT project, anti-epileptic medications, insulin/insulin analogues, anti-asthmatic medications and selective serotonin re-uptake inhibitors and psycholeptics, were followed-up by the relevant EUROmediCAT working groups [30]. The remaining signals with a FDR of 50%, were investigated in greater detail and reported in an accompanying paper [22].

Evaluation of the signal detection system. Evaluation of the signal detection system was performed by comparing the signals obtained with associations reported in a literature review by van Gelder *et al.* [21] from case–control studies. Associations confirmed in at least two studies and not refuted in studies involving ATC-4 or ATC-5 medication exposures were considered. Eight medication–CA combinations were used to evaluate the system including four for valproic acid and one each for naproxen, oxprenolol, phenytoin and progesterone [21]. While this list is not comprehensive it provides an objective set of known associations against which to evaluate the signal detection system.

Results

Signal detection dataset

A total of 30 513 foetuses with medication exposures and CAs (excluding genetic conditions) born from 1995–2011 were extracted from the EUROmediCAT central database. A total of 898 foetuses with isolated congenital dislocation of the hip, 1288 foetuses with no medication exposures recorded at ATC-4 or ATC-5 level and 448 foetuses with only topical medication exposures were excluded, leaving 27 879 foetuses with valid medication exposures (Table 1).

A further 10 655 foetuses with exposures of unknown timing and 2274 foetuses with exposures outside the first trimester were excluded, for a remaining dataset of 14 950 foetuses. In Poland and Wielkopolska over 84% foetuses were excluded. However the distribution of CA types was similar between retained and excluded foetuses, suggesting no bias in the selection of remaining foetuses. Amongst the 14 950 foetuses, 23 101 first trimester, ATC-coded non-topical medication exposures were recorded, for an average of 1.55 first trimester medication exposures per foetus (Table 1). The most common medication exposures were genitourinary system medications and sex hormones (n = 5256), nervous system medications (n = 4181) and anti-infectives for systemic use (n = 3847) (Table 2). Approximately 3% of the ATC-codes were coded as ATC-4 codes (n = 668), rather than ATC-5 codes (n = 22 433).

A total of 410 ATC-5 codes, 223 ATC-4 codes and 60 substances with multiple ATC-codes in the data were eligible as exposure for analysis as at least three foetuses were exposed to each during the first trimester of pregnancy.

Among the 14 950 malformed infants, the most common anomalies were congenital heart defects (n = 5187), ventricular septal defects (n = 2563) and atrial septal defect (n = 1328) (Table 3). Infants can have multiple anomalies and therefore be included in more than one CA subgroup.

Statistical analysis

Forty-thousand three hundred and eighty-five Fisher's exact tests were performed and a total of 1177 medication-CA combinations were significant at the 5% level (Figure 1).

Figures 2 and 3 are 'smile plots'[30] in which the OR is plotted on the x-axis against the Fisher's exact P value on the y-axis for each medication anomaly combination. The red lines correspond to the corrected overall critical P values at different FDRs and datapoints above it correspond to rejected null hypotheses. This allows for instant visualization of both statistical significance and practical implication of test results. Using a FDR of 50%, a total of 30 exposure–CA combinations were considered signals for the ATC-4 level and aggregate medication group exposures (cut off $P = 0.0011$, Figure 2) and 24 exposure–CA combinations were considered signals for the ATC-5 level, which included chemical compounds with multiple ATC-codes (cut-off P = 0.0004, Figure 3). Of the 54 remaining exposure–CA combinations, two were excluded as there were less than three exposed foetuses (Figure 1).

Signal selection process

A signal selection process was performed to eliminate duplicate signals. Seven exposure–CA combinations based on an ATC-4 code were discarded as a more specific ATC-5 code was associated with the same CA. Three exposure–CA combinations involving an aggregate CA (e.g. neural tube defects) and a specific CA (e.g. spina bifida) were associated with the same medication. The associations with the aggregate anomaly were discarded. Finally, three of the remaining 42 exposure–CA combinations were protective and therefore discarded. A total of 39 exposure–CA combinations were considered signals.

Table 2

Number of foetuses with a non-chromosomal congenital anomaly according to exposure to first trimester medication (n = 14 950)

Medication group	ATC-code	Number[a]	Percentage of malformed foetuses exposed (%)
Alimentary tract and metabolism	**A**	**2599**	**17.4**
Antacids and medications for peptic ulcer	A02	730	4.9
Medications for functional gastrointestinal disorders	A03	755	5.1
Antidiarrheals, intestinal anti-inflammatory/anti-infective agents	A07	158	1.1
Antiobesity preparations, excluding diet products	A08	27	0.2
Medications for diabetes	A10	606	4.1
Blood and blood forming organs	**B**	**696**	**4.7**
Antithrombotic agents	B01	637	4.3
Cardiovascular system	**C**	**1042**	**7.0**
Vasoprotectives	C05	158	1.1
β-adrenoceptor blockers	C07	259	1.7
Agents acting on the renin-angiotensin system	C09	55	0.4
Dermatologicals	**D**	**889**	**6.0**
Antifungals for dermatological use	D01	188	1.3
Antibiotics and chemotherapy for dermatological use	D06	131	0.9
Dermal corticosteroids	D07	362	2.4
Anti-acne preparations	D10	74	0.5
Genitourinary system and sex hormones	**G**	**5256**	**35.2**
Gynaecological anti-infectives	G01	502	3.4
Other gynaecologicals	G02	1031	6.9
Sex hormones	G03	3705	24.8
Systemic hormonal preparations, excluding sex hormones and insulins	**H**	**1753**	**11.7**
Posterior pituitary lobe hormones	H01B	8	0.1
Thyroid therapy	H03	1298	8.7
Anti-infectives for systemic use	**J**	**3847**	**25.7**
Antibacterials for systemic use	J01	3399	22.7
Tetracyclines	J01 A	89	0.6
β-lactam antibacterials, penicillins	J01C	2065	13.8
Sulphonamides and trimethoprim	J01E	111	0.7
Macrolides, lincosamides and streptosamins	J01F	478	3.2
Other antibacterials	J01X	384	2.6
Antineoplastic and immunomodulating agents	**L**	**137**	**0.9**
Musculoskeletal system	**M**	**675**	**4.5**
Anti-inflammatory and antirheumatic products, non-steroids	M01A	605	4.1
Muscle relaxants	M03	35	0.2
Nervous system	**N**	**4181**	**28.0**
Opioids	N02 A	294	2.0
Anti-epileptics	N03	675	4.5
Antipsychotics	N05 A	235	1.6

(*continues*)

Table 2

(Continued)

Medication group	ATC-code	Number[a]	Percentage of malformed foetuses exposed (%)
Anxiolytics, hypnotics and sedatives	N05C	505	3.4
Antidepressants	N06 A	685	4.6
Antiparasitic products, insecticides and repellents	**P**	**113**	**0.8**
Antiprotozoals	P01	96	0.6
Respiratory system	**R**	**3620**	**24.2**
Nasal preparations	R01	430	2.9
Anti-asthmatics	R03	1868	12.5
Cough and cold preparations	R05	221	1.5
Antihistamines	R06	1027	6.9
Sensory organs and various	**S + V**	**539**	**3.6**

[a]The total number of foetuses is not the sum of all the numbers in this column as each foetus can be exposed to more than one medication

Table 3

Congenital anomaly subgroups analysed for the purpose of signal detection

Congenital anomaly subgroup[a]	Number	Percentage of malformed foetuses affected (%)
Neural tube defects	563	3.77
Anencephalus	163	1.09
Encephalocele	85	0.57
Spina bifida	315	2.11
Hydrocephaly	305	2.04
Microcephaly	118	0.79
Arhinencephaly/holoprosencephaly	44	0.29
Anophthalmos/microphthalmos	74	0.49
Congenital cataract	86	0.58
Congenital glaucoma	27	0.18
Anotia	22	0.15
Congenital heart defects	5187	34.70
Severe congenital heart defects	1208	8.08
Common arterial truncus	37	0.25
Transposition of great vessels	236	1.58
Single ventricle	56	0.37
Ventricular septal defect	2563	17.14
Atrial septal defect	1328	8.88
Atrioventricular septal defect	141	0.94
Tetralogy of Fallot	208	1.39
Tricuspid atresia and stenosis	58	0.39
Ebstein's anomaly	31	0.21
Pulmonary valve stenosis	305	2.04

(*continues*)

Table 3

(Continued)

Congenital anomaly subgroup[a]	Number	Percentage of malformed foetuses affected (%)
Pulmonary valve atresia	70	0.47
Aortic valve atresia/stenosis	115	0.77
Hypoplastic right heart	18	0.12
Coarctation of aorta	220	1.47
Total anomalous pulmonary venous return	29	0.19
PDA in term (> = 37 weeks)	253	1.69
Choanal atresia	35	0.23
Cleft lip ± palate	703	4.70
Cleft palate	486	3.25
Oesophageal atresia	181	1.21
Duodenal atresia or stenosis	54	0.36
Atresia or stenosis of other parts of the small intestine	61	0.41
Ano-rectal atresia and stenosis	224	1.50
Hirschprung's disease	61	0.41
Atresia of the bile ducts	17	0.11
Annular pancreas	15	0.10
Diaphragmatic hernia	182	1.22
Gastroschisis	166	1.11
Omphalocele	137	0.92
Bilateral renal agenesis	74	0.49
Renal dysplasia	245	1.64
Congenital hydronephrosis	792	5.30
Bladder exstrophy and/or epispadia	44	0.29
Posterior urethral valve and/or prune belly	69	0.46
Hypospadias	1290	8.63
Limb reduction	426	2.85
Complete absence of limb	13	0.09
Club foot	863	5.77
Polydactyly	610	4.08
Syndactyly	378	2.53
Craniosynostosis	130	0.87
Congenital construction bands	31	0.21
Conjoined twins	5	0.03
Laterality[b]	93	0.62
Situs inversus	49	0.33
Neural crest defects[c]	816	5.46

[a]Except for the two newly formed subgroups of laterality and neural crest, congenital anomaly subgroups were based on EUROCAT defined subgroups [24]
[b]The newly formed laterality subgroup consists of atrial isomerism, dextrocardia, situs inversus, broncho-pulmonary isomerism and asplenia and polysplenia
[c]The newly formed neural crest defects subgroup consists of coloboma, Hirschprung's disease, Pierre Robin sequence, Goldenhar, double outlet left/right ventricle, truncus arteriosus, interrupted aorta, double aortic arch, transposition of great arteries, Fallot's tetralogy and Arnold Chiari I/II malformations

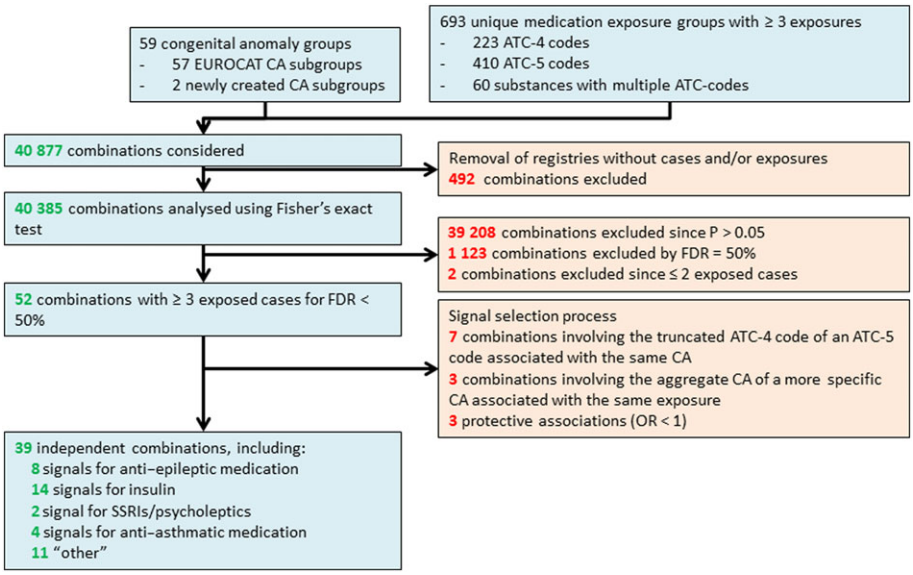

Figure 1

Flow chart of congenital anomaly medication exposure combinations

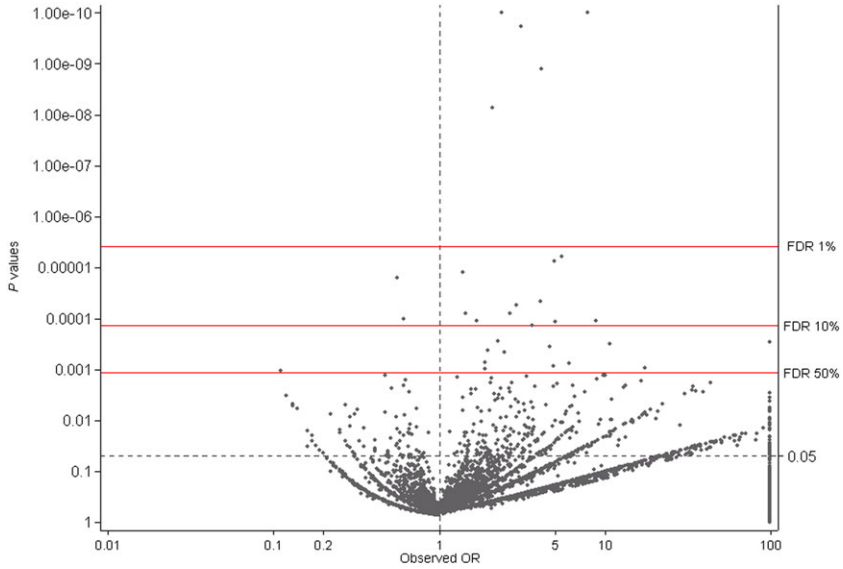

Figure 2

Smile plot for the Simes multiple testing procedure for ATC-4 codes plus aggregate groups (n = 13 065 exposure–CA combinations; odds ratio truncated at 100 and P value truncated at 1.00E – 10)

Evaluation of the signal detection system

Of the 39 exposure–CA combinations, eight related to anti-epileptics (of which two were valproic acid) which are well established to be teratogenic and 14 related to insulin, a marker of maternal diabetes which is well established to be teratogenic. Signals relating to the less well established effects of selective serotonin re-uptake inhibitors/psycholeptics and antiasthmatics were evaluated in parallel analyses of the database [31]. Other signals are discussed in an accompanying paper [22].

Out of the eight van Gelder *et al*. [21] signals used to evaluate the system our methodology identified two of the valproic acid signals with a FDR of 50% (Table 4). The other two valproic acid signals were statistically significant at the 5% level. However their FDR values were above the 50% cut-off. Naproxen and progesterone had an increased OR but were not statistically significant although there were related signals for sex hormone medications. Oxprenolol and phenytoin had too few exposures to be analyzed (nought exposures and 18 exposures with one exposed case respectively).

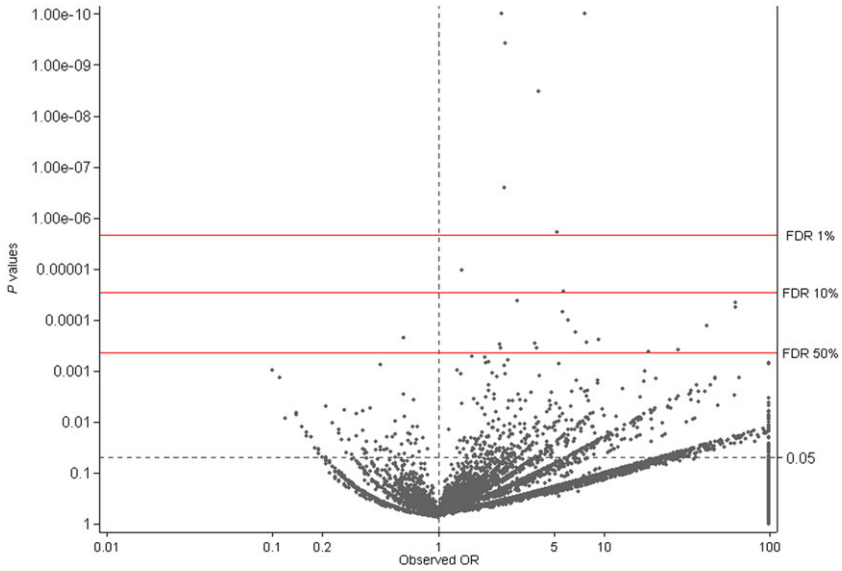

Figure 3

Smile plot for the Simes multiple testing procedure for ATC-5 codes plus substances with multiple ATC-codes (n = 27 320 exposure–CA combinations; odds ratio truncated at 100 and P value truncated at 1.00E – 10)

Table 4

Evaluation of selected known medication–congenital anomaly associations identified by van Gelder *et al.* [21]

ATC	Medication	Congenital anomaly	Exposed foetuses in EUROmediCAT	Exposed foetuses with specified congenital anomaly in EUROmediCAT	OR (95% CI)	*P*	FDR (%)
G02CC02	Naproxen	Cleft lip ± palate	28	3	2.58 (0.50, 8.50)	0.13	
C07AA02	Oxprenolol	Cleft lip ± palate	0	0	NA	NA	
N03AB02	Phenytoin	Cleft lip ± palate	18	1	NA	NA	
G03DA04	Progesterone	Hypospadias	1074	103	1.13 (0.91, 1.40)	0.14	
N03AG01	Valproic acid	ASD	223	32	1.74 (1.15, 2.56)	0.004	>50
N03AG01	Valproic acid	Cleft palate	223	20	3.01 (1.78, 4.82)	0.00004	<50
N03AG01	Valproic acid	Craniosynostosis	222[a]	5	2.75 (0.87, 6.70)	0.04	>50
N03AG01	Valproic acid	Spina bifida	223	28	7.70 (4.89, 11.75)	< 0.0001	<1

[a]A registry without any cases of craniosynostosis, but with a single valproic acid exposure was removed from analysis.Hence only 222 rather than 223 exposures were analyzed

Discussion

This study showed a systematic process of selecting 39 signals out of a starting total of over 40 000 exposure–CA combinations, by controlling the FDR. Although signal detection using databases containing exclusively medication exposed cases of CAs is not novel (for example the SAFE-Med project [20]), controlling the FDR is new in this context. The entire selection process can be automated, but signals generated require detailed follow-up [22].

Statistical considerations

Due to the large number of analyses performed and therefore the large number of false positives expected when using the conventional 5% level of significance, multiple testing procedures were applied. Multiple testing procedures can be designed to control the familywise error rate or the FDR. Procedures controlling the familywise error rate, such as Bonferroni & Sidak, control the chance of even a single false positive result at a 5% level of significance [32]. Controlling the familywise error rate comes at a great loss of statistical

power as the number of hypotheses in the family increases. An alternative was proposed by Benjamini & Hochberg [28] in controlling the FDR, rather than the familywise error rate. A number of alternatives are available for controlling the FDR [33], of which the Simes procedure is more inclusive than others such as Yekutieli. It was decided to use the Simes method as we planned to follow-up the results and so missing potential associations was more of a concern than identifying potential false positives.

Analyzing associations for more common anomalies (for example all heart anomalies) and common medications (for example aspirin) separately from the rarer anomalies and medication exposures was considered, as the more common associations have a much greater power for being detected. However analyzing the more common associations separately from the rarer associations could result in a biased set of data as a biased sample is removed from the data.

Three protective associations were identified: atrial septal defect in combination with G02CA (sympathomimetics, labour repressants), N02BE (anilides) and N02BE01 (paracetamol). The restrictions in this study of only including malformed foetuses and foetuses exposed to at least one medication means that the power to detect any protective associations is extremely limited. Protective associations are likely to arise either due to chance (with a FDR of 50%, 50% of observed associations are likely to have arisen by chance) or due to bias arising from the study design. For instance there is a strong association between valproic acid and spina bifida. If the association between valproic acid and a different anomaly is examined, the controls will include a large number of anomalies associated with valproic acid and the protective effect observed may actually be due to the large numbers of controls being associated with valproic acid. The fact that only three positive associations were identified demonstrates that it is not likely to be a problem in this population-based surveillance system.

Non-chromosomal controls were used rather than chromosomal controls. Although chromosomal anomalies are not caused by medication exposure during pregnancy and are well suited as controls, they are rarer than non-chromosomal anomalies. Chromosomal anomalies make up approximately 15% of all CAs [19], thereby reducing the power to detect significant results. In addition statistical adjustment for maternal age is required when using chromosomal controls due to the link with maternal age and chromosomal anomalies and the frequent link between maternal age and medication exposure.

The multiple testing procedure was applied separately to aggregate medication groups including the ATC-4 codes, ATC-5 codes and chemical compound names. These sets of results were considered separately due to the overlap in cases between ATC-4 and ATC-5 codes.

A method based solely on automated statistical analyses is suitable for the first stage in signal detection in order to reduce the large number of 'statistically significant' associations to those that are more likely to be true associations (judged by the FDR). The selection of the cut-off value of the FDR is a trade-off between the proportion of false positives and false negatives. Too high a FDR will result in a large amount of work investigating associations and potentially causing unjustified anxiety amongst pregnant women. Too low a FDR will result in a failure to detect teratogenic medications in a timely manner. This means that the FDR cut-off should be re-evaluated regularly and depends on human resources available for follow-up of signals. However, this signal detection process is not the only analysis of the data. Each medication class is also separately analyzed in turn in detailed analyses emphasising hypothesis testing, limiting the implications of setting the FDR criterion too high.

The second stage in signal detection considers factors such as the quality of the data (including registry-effects), concomitant medication exposures and consistency with the literature, which are discussed in the accompanying paper [22], as well as biological plausibility [34].

Data in the signal detection dataset were supplied by a number of registries and differences in medication usage, coding and ascertainment between registries can lead to bias. For example certain medications are used only by women from specific registers. If these registers also have a higher prevalence of specific anomalies than other registers then it will appear that the medications are linked to those anomalies. It would be advantageous to adjust for registry and other potential confounders such as maternal age prior to performing the multiple comparison procedure. However the large number of analyses involving extremely small numbers of exposures prohibits accurate adjustment for confounders. Therefore the multiple comparison procedure was performed on unadjusted results. Heterogeneity between registries and registry-specific effects are considered in the accompanying paper examining the identified signals in detail [22].

Strengths and weaknesses

EUROmediCAT is an international population-based database that contains both detailed information on medications taken during the first trimester of pregnancy and detailed coding of all CAs. EUROmediCAT has advantages over other hypotheses generating databases, such a VigiBase™, EudraVigilance and FAERS that rely on spontaneous reports. These other databases have more limited coding for CAs, cannot distinguish between chromosomal and non-chromosomal CAs, are not able to identify first trimester exposures and are vulnerable to disproportional spontaneous reporting, for example caused by media attention.

A potential weakness of EUROmediCAT is lack of information on duration, dose of medication exposure and adherence. A study using additional data sources found that there was a high degree of agreement between the medication actually used and that recorded in one EUROmediCAT registry [35]. There is considerable heterogeneity in exposure between the registries, which is likely to reflect true registry differences as well as differences in reporting. The data cleaning attempted to ensure only actual exposures are included, at the risk of missing some exposures, which will have the effect of reducing the power to detect true associations, due to fewer cases, but should not produce biased estimates. A large proportion (84%) of the Polish CA registrations were excluded following data cleaning since we could not be sure that exposures occurred in the first trimester of pregnancy. Similar distributions of CAs were observed in excluded and retained registrations, but bias could have occurred.

The non-chromosomal malformed controls utilized could lead to potential underestimation of effect size when part of the control group is related to the medication investigated. This process is also known as 'masking'. It could be considered to exclude known associations in future EUROmediCAT signal detection algorithms once a clear definition can be provided for 'known associations'. Additionally, artefacts such as confounding by indication (for example diabetes in the mother is a risk factor of congenital heart disease in the foetus, with the consequence that antidiabetic medication exposure is expected to be associated with congenital heart disease) can also lead to masking.

Evaluation of the signal detection system

The detection of eight associations involving anti-epileptics (including two for valproic acid) and 14 involving insulin (a marker of maternal diabetes which is well established to be teratogenic) provides evidence that this methodology will detect some known teratogens. However it does not provide any measure of how many teratogens are not detected and the reasons for the lack of detection. Theoretical considerations would lead us to predict that teratogens will be more likely to be missed if exposure levels are very low, or the relative risk low, or they are non-specific in the type of congenital anomalies caused. To look at this quantitatively a definitive set of medication anomaly associations should be identified and the proportion detected calculated. However it is difficult to identify 'known' medication anomaly associations. A recent comprehensive review highlighted the lack of evidence available for the teratogenic risk of medications commonly used in pregnancy [21]. Only eight associations with considerable evidence were able to be selected objectively to evaluate the system. Two of four concerning valproic acid were identified as signals by the system using 50% FDR. Two signals (for phenytoin and oxprenolol) could not be tested in the EUROmediCAT data as there were insufficient exposed cases. This is likely to be due to the fact that the teratogenic effect of these medications is known and so their use in pregnant women has been reduced and it is therefore not a weakness of the signal detection system. This may also explain why naproxen (with only 28 exposures) was not statistically significant in our study. Two signals with valproic acid were not identified as signals despite having a considerable number of exposures in the database, but other anomalies more strongly associated with valproic acid were picked up as signals. Progesterone was not picked up as a signal although another sex hormone signal was identified. This gives an indication of the potential weakness of the system. Not all associations will be detected and it may be that the FDR of 50% should be even higher in future analysis in order to identify more signals. This would need to be balanced against the time and resources needed to explore the new signals identified [22]. No signal detection system will identify all associations using this specific methodology. More targeted studies can be performed on specific anomaly medication combinations with prior hypotheses. Hence the other EUROmediCAT workpackages [36] analyzing individual medication groups using different methodology and different data sets were important to identify other associations.

In conclusion data on medication exposure in European CA registers can be analyzed systematically using the FDR to determine a manageable set of potential associations for further investigation. This methodology was able to identify some known teratogens, but not all. The value of the new associations identified is determined in a separate paper [22]. The signals identified could be valuable in prioritizing efforts for conducting hypothesis-driven studies.

Competing Interests

All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare the study was funded by the European Commission under the 7th Framework Program (grant agreement HEALTH-F5-2011-260598). The institutions of HD, ML, JM, EG, MCA, IB, MG, BK, AL-B, VN, AJN, MO'M, AP, DT and AW received funding from Glaxo Smith Kline (GSK) outside the submitted work in the previous 3 years. There are no other relationships or activities that could appear to have influenced the submitted work.

Contributors

HD, JG, JM, JML, LJB and ML planned and designed the study. JM and JML drafted the manuscript and performed the analyses. ALB, AN, AP, AW, DT, EG, HW, IB, KK, MB, MCA, MG, MO, NL and VN contributed data as registry staff. All authors reviewed the manuscript and approved of the final version.

References

1 Botting J. The history of thalidomide. Drug News Perspect 2002; 15: 604–11.

2 Jentink J, Loane MA, Dolk H, Barisic I, Garne E, Morris JK, *et al*. Valproic acid monotherapy in pregnancy and major congenital malformations. N Engl J Med 2010; 362: 2185–93.

3 Hobbs CA, Chowdhury S, Cleves MA, Erickson S, MacLeod SL, Shaw GM, *et al*. Genetic epidemiology and nonsyndromic structural birth defects: from candidate genes to epigenetics. JAMA Pediatr 2014; 168: 371–7.

4 Nau H. Teratogenicity of isotretinoin revisited: species variation and the role of all-trans-retinoic acid. J Am Acad Dermatol US 2001; 45: S183–S187.

5 Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. Am J Med Genet C Semin Med Genet 2011; 157C: 175–82.

6 Bakker MK, Jentink J, Vroom F, Van Den Berg PB, De Walle HE, De Jong-Van Den Berg LT. Drug prescription patterns before, during and after pregnancy for chronic, occasional and pregnancy-related drugs in the Netherlands. BJOG 2006; 113: 559–68.

7 Engeland A, Bramness JG, Daltveit AK, Ronning M, Skurtveit S, Furu K. Prescription drug use among fathers and mothers before and during pregnancy. A population-based cohort study of

106,000 pregnancies in Norway 2004-2006. Br J Clin Pharmacol 2008; 65: 653–60.

8 Singh S, Sedgh G, Hussain R. Unintended pregnancy: worldwide levels, trends, and outcomes. Stud Fam Plann 2010; 41: 241–50.

9 Harmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. Eur J Clin Pharmacol 2008; 64: 743–52.

10 Mcbride WG. Thalidomide and congenital abnormalities. Lancet 1961; 278: 1358.

11 Lindquist M. VigiBase, the WHO global ICSR database system: Basic facts. Drug Inf J 2008; 42: 409–19.

12 Alvarez Y, Hidalgo A, Maignen F, Slattery J. Validation of statistical signal detection procedures in eudravigilance post-authorization data: a retrospective evaluation of the potential for earlier signalling. Drug Saf 2010; 33: 475–87.

13 Sakaeda T, Tamon A, Kadoyama K, Okuno Y. Data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci 2013; 10: 796–803.

14 Sharrar RG, Dieck GS. Monitoring product safety in the postmarketing environment. Ther Adv Drug Saf 2013; 4: 211–19.

15 Suling M, Pigeot I. Signal detection and monitoring based on longitudinal healthcare data. Pharmaceutics 2012; 4: 607–40.

16 Edwards IR, Biriell C. Harmonisation in pharmacovigilance. Drug Saf 1994; 10: 93–102.

17 Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systematic review. Drug Saf 2006; 29: 385–96.

18 Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Perucca E, *et al*. EURAP: an international registry of antiepileptic drugs and pregnancy. Epilepsia 2004; 45: 1463–4.

19 EUROCAT Central Registry. 2012. EUROCAT prevalence tables [online]. Available at http://www.eurocat-network.eu/accessprevalencedata/prevalencetables (accessed 24 January 2014).

20 Lisi A, Botto LD, Robert-Gnansia E, Castilla EE, Bakker MK, Bianca S, *et al*. Surveillance of adverse fetal effects of medications (SAFE-Med): findings from the international Clearinghouse of birth defects surveillance and research. Reprod Toxicol 2010; 29: 433–42.

21 van Gelder MM, de Jong-van den Berg LT, Roeleveld N. Drugs associated with teratogenic mechanisms. Part II: a literature review of the evidence on human risks. Hum Reprod 2014; 29: 168–83.

22 Given JE, Loane M, Luteijn JM, Morris JK, de Jong-van den Berg L, Garne E, *et al*. EUROmediCAT signal detection: an evaluation of selected congenital anomaly-medication associations. Br J Clin Pharmacol 2016; 82: 1094–109.

23 Boyd PA, Haeusler M, Barisic I, Loane M, Garne E, Dolk H. Paper 1: The EUROCAT network--organization and processes. Birth Defects Res A Clin Mol Teratol 2011; 91 (Suppl 1): S2–15.

24 Loane M, Dolk H, Garne E, Greenlees R, EUROCAT Working Group. Paper 3: EUROCAT data quality indicators for population-based registries of congenital anomalies. Birth Defects Res A Clin Mol Teratol 91: S23–S30.

25 Registry EC. 2005. EUROCAT Guide 1.3 and reference documents [online]. Available at http://www.eurocat-network.eu/content/EUROCAT-guide-1.3.pdf (accessed 06/07 2012).

26 EUROCAT. 2014. Special Report: Sources of Information on Medication Use in Pregnancy [online]. Available at http://www.eurocat-network.eu/content/Special-Report-Medication-Use-In-Pregnancy.pdf (accessed 17 January 2016).

27 World Health Organization. 2015. ATC alterations from 1982-2015 [online]. Available at http://www.whocc.no/atc_ddd_alterations__cumulative/atc_alterations/ (accessed 13 March 2015).

28 Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J Roy Stat Soc B 1995; 57: 289–300.

29 Stata statistical software. StataCorp, Release 12 ed. College station, TX: StataCorp LP, 2011.

30 Newson R, The ALSPAC Study Team. Multiple-test procedures and smile plots. Stata J 2003; 3: 109–32.

31 EUROmediCAT. 2015. Research Programme [online]. Available at http://euromedicat.eu/fp7researchprogramme (accessed 30 September 2015).

32 Šidák Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. J Am Stat Assoc 1967; 62: 626–33.

33 Benjamini Y, Yuketieli D. The Control of the False Discovery Rate in Multiple Testing under Dependency. Ann Stat 2001; 29: 1165–88.

34 Meyboom RH, Lindquist M, Egberts AC, Edwards IR. Signal selection and follow-up in pharmacovigilance. Drug Saf 2002; 25: 459–65.

35 de Jong L, de Walle HEK, de Jong-van den Berg LTW, van Langen IM, Bakker MK. Actual use of medications prescribed during pregnancy: A cross-sectional study using data from a population-based congenital anomaly registry. Drug Saf 2015; 38: 737–47.

36 EUROmediCAT. 2015. EUROmediCAT publications [online]. Available at http://www.euromedicat.eu/publicationsand presentations/publications (last accessed 26 July 2016).