Received: 23 November 2017 | Revised: 20 March 2018 | Accepted: 12 April 2018

DOI: 10.1002/pds.4556

ORIGINAL REPORT

WILEY

Development of an algorithm to identify pregnancy episodes and related outcomes in health care claims databases: An application to antiepileptic drug use in 4.9 million pregnant women in France

Pierre-Olivier Blotière[1,2] | Alain Weill[1] | Marie Dalichampt[1] | Cécile Billionnet[1] | Myriam Mezzarobba[1] | Fanny Raguideau[3] | Rosemary Dray-Spira[3] | Mahmoud Zureik[3] | Joël Coste[1] | François Alla[2,4]

[1] Department of Studies in Public Health, French National Health Insurance (CNAMTS), Paris, France

[2] Université de Lorraine, université Paris-Descartes, Apemac, EA 4360, Nancy, France

[3] Department of Epidemiology of Health Products, The French National Agency for Medicines and Health Products Safety, Saint-Denis, France

[4] Université Bordeaux et CHU Bordeaux, CIC-EC 1401, Inserm U1219, Bordeaux, France

Correspondence
P.-O. Blotière, French National Health Insurance (CNAMTS), 50 Avenue du Professeur André Lemierre, 75986 Paris Cedex 20, France.
Email: pierre-olivier.blotiere@cnamts.fr

Abstract

Purpose: Access to claims databases provides an opportunity to study medication use and safety during pregnancy. We developed an algorithm to identify pregnancy episodes in the French health care databases and applied it to study antiepileptic drug (AED) use during pregnancy between 2007 and 2014.

Methods: The algorithm searched the French health care databases for discharge diagnoses and medical procedures indicative of completion of a pregnancy. To differentiate claims associated with separate pregnancies, an interval of at least 28 weeks was required between 2 consecutive pregnancies resulting in a birth and 6 weeks for terminations of pregnancy. Pregnancy outcomes were categorized into live births, stillbirths, elective abortions, therapeutic abortions, spontaneous abortions, and ectopic pregnancies. Outcome dates and gestational ages were used to calculate pregnancy start dates.

Results: According to our algorithm, live birth was the most common pregnancy outcome (73.9%), followed by elective abortion (17.2%), spontaneous abortion (4.2%), ectopic pregnancy (1.1%), therapeutic abortion (1.0%), and stillbirth (0.4%). These results were globally consistent with French official data. Among 7 559 701 pregnancies starting between 2007 and 2014, corresponding to 4 900 139 women, 6.7 per 1000 pregnancies were exposed to an AED. The number of pregnancies exposed to older AEDs, comprising the most teratogenic AEDs, decreased throughout the study period (−69.4%), while the use of newer AEDs increased (+73.4%).

Conclusions: We have developed an algorithm that allows identification of a large number of pregnancies and all types of pregnancy outcomes. Pregnancy outcome and start dates were accurately identified, and maternal data could be linked to neonatal data.

The illustration part of this work has been previously presented at the 33rd International Conference on Pharmacoepidemiology and Therapeutic Risk Management in Montreal, Canada (August 2017).

Joël Coste and François Alla are co-last authors.

KEYWORDS

algorithm, antiepileptic drugs, claims data, French health care databases, pharmacoepidemiology, pregnancy

1 | INTRODUCTION

The exclusion of pregnant women from clinical trials results in a lack of information about the effects of medication use during pregnancy on maternal and fetal health. Access to claims databases therefore provides an opportunity to study medication use and safety during pregnancy. These databases are particularly useful, as these studies, often concerning rare exposures and rare birth outcomes, require large data sources. Routine post-marketing research is also possible, since information on a wide range of outcomes and all prescription drugs prescribed or dispensed during pregnancy is available. Ascertainment of medication use is based on pharmacy claims data and is independent of maternal or infant outcomes, which avoids parental recall bias.[1]

However, claims databases are usually not built for research purposes and algorithmic approaches must be developed in order to identify pregnancy episodes and related outcomes.[2] A few articles have been specifically devoted to this challenge, using the Clinical Practice Research Datalink[2,3] or North American claims databases.[1,4-8] In France, despite growing interest in pregnancy research using claims databases,[9-16] no algorithm has yet been published. However, some studies have explicitly reported the discharge diagnoses or medical procedures used to identify pregnancies.[9-11,13,16] Most studies were restricted to births, and only a few studies included abortion and ectopic pregnancies.[9,12,14,16]

The primary objective of this study was therefore to develop an algorithm to identify pregnancy episodes and related outcomes using the French health care databases, which covers 99% of the 67 million inhabitants in France.[17] The secondary objective was to apply this algorithm to the analysis of antiepileptic drug (AED) use during pregnancy in France between 2007 and 2014. Studying AED use during pregnancy is of particular interest, as prenatal exposure to some older AEDs has been found to be associated with increased risks of major congenital malformations,[18] and prenatal exposure to valproic acid has been found to be associated with an increased risk of autism spectrum disorder.[19]

KEY POINTS

- Access to claims databases provides an opportunity to study medication use and safety during pregnancy.
- Few articles have been specifically devoted to identification of pregnancies in claims databases, and no algorithms based on the French health care databases have been published.
- We have developed an algorithm that captured all types of pregnancy outcome and accurately identified pregnancy episodes.
- Among 7 559 701 pregnancies starting between 2007 and 2014, 6.7 per 1000 pregnancies were exposed to an AED.

2 | METHODS

2.1 | Data source

This study was conducted using the French national health insurance information system (SNIIRAM), which consists of 2 French nationwide datasets linked by a unique patient identifier: the French national health insurance database (DCIR) and the French hospital discharge database (PMSI). French national health insurance covers the entire French population and is divided into several specific schemes, including the general scheme for salaried workers (87% of the population), the self-employed workers scheme (6%), the farmers scheme (5%), and other additional schemes covering the remaining 2%.[20]

The DCIR database contains all individualized and anonymous health care claims reimbursed by French National Health Insurance. These claims include, in particular, dispensed drugs and medical procedures. The DCIR database also collects patient data such as age, gender, and eligibility for 100% health insurance coverage for serious and costly long-term diseases (LTD) coded according to the International Classification of Diseases, 10th Revision (ICD-10), but does not contain other outpatient medical indications. The PMSI database provides detailed medical information on all admissions in public and private hospitals in France, including discharge diagnosis ICD-10 codes and medical procedures coded according to the French medical classification for clinical procedures (CCAM).

2.2 | Algorithm

Pregnancies were identified on the basis of their outcome from the DCIR database for outpatient medical abortions and from the PMSI database for all other pregnancy outcomes. The algorithm searched the databases for discharge diagnoses and medical procedures indicative of completion of a pregnancy coded between 2007 and 2015. Diagnoses and procedures are presented in Table 1. All records associated with unknown women identifiers, which could not be linked to any other data, were discarded. A flowchart describing the algorithm is available in Supporting Information (supplementary figure 1).

In the first step, all records of codes indicative of completion of a pregnancy in the PMSI database were grouped into 2 categories: (1) births (end of pregnancy ≥22 weeks after the last menstrual period [LMP]); (2) any terminations of pregnancy <22 weeks after the LMP. Codes representing the same pregnancy were then removed: for each woman, duplicate records were addressed separately in the 2 groups by choosing the last code as the pregnancy outcome within a predetermined time-frame. A 28-week span was used for births, and

TABLE 1 Data used to identify pregnancy outcomes in the SNIIRAM databases

Live births	Associated diagnoses Z37, Z3900[a] or principal diagnoses O80, O81, O82, O83, O84 OR delivery procedure[b] WITHOUT diagnoses indicative of stillbirth or therapeutic abortions ≥22 weeks after the LMP
Stillbirths	Associated diagnoses Z37.1, Z37.3, Z37.4, Z37.6, Z37.7 WITHOUT principal diagnosis O35 before March 2011 Associated diagnoses Z37.10, Z37.30, Z37.40, Z37.60, Z37.70[c] after March 2011
Elective abortions	
Inpatient elective abortions	Principal diagnoses O04, O05, O06, O07 AND procedure indicative of inpatient abortion[d] AND associated diagnosis Z640
Outpatient medical abortions	Procedure indicative of outpatient medical abortion[e]
Therapeutic abortions	
<22 weeks after the LMP	Principal diagnoses O04, O05, O06, O07 AND procedure indicative of inpatient abortion[d] WITHOUT associated diagnosis Z640
≥22 weeks after the LMP	Associated diagnoses Z37.1, Z37.3, Z37.4, Z37.6, Z37.7 AND principal diagnosis O35 before March 2011 Associated diagnoses Z37.11, Z37.31, Z37.41, Z37.61, Z37.71[c] after march 2011
Other abortions	Principal diagnoses O04, O05, O06, O07 WITHOUT procedure indicative of inpatient abortion[d]
Spontaneous abortions	Principal diagnosis O03
Ectopic pregnancies	Principal diagnosis O00 OR procedure indicative of ectopic pregnancy[f]
Others[g]	Principal diagnosis O01, O02

Note:

Stillbirth = death of a fetus with a gestational age ≥ 22 weeks after the LMP or with a birth weight ≥ 500 g.

Spontaneous abortion = death of a fetus with a gestational age < 22 weeks after the LMP and a birth weight < 500 g.

Elective abortion = termination of pregnancy at the woman's request for reasons other than maternal health or fetal disease, possible until 14 weeks after the LMP in France.

[a]Care and examination immediately after delivery outside hospital.

[b]CCAM codes JQGD010, JQGD012, JQGD004, JQGD001, JQGD003, JQGD008, JQGD013, JQGD005, JQGD002, JQGD007, JQGA002, JQGA004, JQGA003, JQGA005.

[c]The extension "0" indicates stillbirth, excluding therapeutic abortion and "1" indicates therapeutic abortion.

[d]CCAM codes JNJD001, JNJD002 (surgical abortion), JNJP001 (medical abortion).

[e]Outpatient procedure codes 2422, 3329 (management of medical abortion), 2415 (mifepristone), 2416 (prostaglandin), available only for the general scheme before 2009.

[f]CCAM codes JJFA001, JJFC001 (Salpingectomy), JJJA002, JJJC002 (fimbrial evacuation), JJLJ001 (*In situ* injection of methotrexate), JJPA001, JJPC001 (salpingostomy), JQGA001 (Removal of abdominal pregnancy more than 13 weeks after the LMP).

[g]Hydatidiform mole or other abnormal products of conception.

a 6-week span was used for terminations of pregnancy, as consecutive pregnancies within these time-frames were deemed implausible.[21]

In the second step, all records of codes related to outpatient medical abortions were identified in the DCIR database, and duplicate records were addressed in a similar way to duplicate records of terminations of pregnancy in the PMSI database. Only outpatient medical abortions performed outside a 6-week span before and after a termination of pregnancy identified in the first step were included.

In the third step, terminations of pregnancy occurring during a pregnancy resulting in a birth were excluded. Terminations of pregnancy occurring during the first 10 weeks after a birth were also excluded.[21]

Pregnancy outcomes were finally categorized into live births, stillbirths, elective abortions (both inpatient and outpatient), therapeutic abortions, spontaneous abortions, ectopic pregnancies, and other outcomes (hydatidiform mole or other abnormal products of conception).

2.2.1 | Pregnancy start date

Pregnancy start dates were calculated from the following:

1. pregnancy outcome dates. Exact admission, discharge, and medical procedure dates have been recorded in the PMSI database since 2009 (supplementary figure 2). When a medical procedure was performed, the exact procedure date was used as the outcome date, which was the case for 97.0% of all births and 91.7% of all inpatient induced abortions. Otherwise, the exact admission date was used. Before 2009, only discharge months were available in the PMSI database, and the outcome date was considered to be the fifteenth day of the discharge month.
2. gestational ages or numbers of days after the LMP. Gestational age has been recorded in the PMSI database since March 2008[22] and exhaustively since March 2010 for all births. It is expressed in completed gestational weeks and has been validated,[23] with a high positive predictive value.[24] For inpatient abortions and other pregnancy outcomes, the number of days after the LMP has been recorded since March 2011[22] and exhaustively since March 2012 (supplementary figure 2). The median gestational ages observed in the PMSI database in 2014

were therefore used to replace missing gestational ages or numbers of days after the LMP according to the type of pregnancy outcome. For this purpose, abortions were further detailed according to trimester and method (supplementary table 1).

This pregnancy start date was compared with an estimated conception date recorded in the DCIR database independently of the information available in the PMSI database. This estimated conception date is reported only for women entitled to maternity leave, which excludes all pregnancies ending before 22 weeks and self-employed workers, regardless of the pregnancy outcome.

2.2.2 | Comparison with official national data

The number of pregnancies identified by the algorithm was compared with official national data for 2014. Official data on live births are published by the National Institute of Statistics and Economic Studies (INSEE), which records all births occurring in France. Data on therapeutic abortions are published by the French Biomedicine Agency and correspond to the number of authorized abortions and not the total number of abortions actually performed. Official data on elective abortions, corresponding to the number of abortion procedures and not the number of distinct pregnancies, are based on the French health care databases, not allowing any valid comparisons.

2.2.3 | Linkage between maternal and neonatal data

Linkage between maternal and neonatal data has been possible in the PMSI database since 2011 by means of a common identifier shared by the mother and her child and present in both the delivery stay and the birth stay. As a birth stay is coded in the PMSI database only for children with a gestational age ≥ 22 weeks after the LMP, this linkage is possible for live births, stillbirths, and therapeutic abortions after 22 weeks.

2.3 | Antiepileptic drug use

As women may have multiple pregnancies during the study period, the unit of analysis was a pregnancy. All pregnancies starting between 2007 and 2014, regardless of the outcome, were eligible for inclusion. The mother had to have continuous health insurance enrolment for a 1-year period before pregnancy. The study was based on the national health insurance general scheme to ensure complete availability of data throughout the study period.

AEDs were defined according to the World Health Organization Anatomical Therapeutic Chemical classification (supplementary table 2). AEDs marketed before the early 1990s are traditionally referred to as "older" AED, whilst drugs that were introduced later are referred to as "newer" AED.[25] Women were considered to be exposed during the 30 days following dispensing.

Prevalence of AED use during pregnancy was assessed between 2007 and 2014. Prevalence was defined as the number of pregnancies exposed to AEDs per 1000 pregnancies. Prevalence rates were calculated overall, by drug group (older versus newer AEDs) and by Anatomical Therapeutic Chemical classes. Drug use was also described by trimester of pregnancy: day 0 to day 90 (first trimester), day 91 to day 181 (second trimester), and day 182 until delivery (third trimester). If a period of exposure began in a given trimester and carried over into the subsequent trimester, both trimesters were considered to be exposed. Trends were also investigated, especially for pregnant women with epilepsy, identified with LTD codes G40 and G41. All epileptic women without epilepsy recorded as an LTD were missed.

A sensitivity analysis was conducted to account for uncertainty in estimating the time period during which a woman was pregnant: a lower limit for prevalence rates was calculated using the 5th percentile of gestational age instead of the median when gestational age was missing and an upper limit was calculated using the 95th percentile of gestational age.

3 | RESULTS

3.1 | Algorithm

The algorithm identified 6 230 200 women who had 9 647 843 pregnancies between 2007 and 2015 (Table 2). Mean age at the end of pregnancy was 29.5 years. Live birth was the most common pregnancy outcome (73.9%), followed by elective abortion (17.2%), spontaneous abortion (4.2%), ectopic pregnancy (1.1%), therapeutic abortion (1.0%), and stillbirth (0.4%). From 2009 onwards, the estimated conception date available for women entitled to maternity leave was

TABLE 2 Distribution of pregnancy episodes by maternal age, type of outcome and twin pregnancies over the 2007 to 2015 study period

	n	%
Maternal age at the end of pregnancy (years)		
Mean (±STD)	29.5 (± 5.9)	
12–19	425 596	4.4%
20–29	4 460 709	46.2%
30–39	4 283 905	44.4%
40–49	471 374	4.9%
50–59	2307	0.0%
Unknown	3952	0.0%
Pregnancy outcome		
Live births	7 126 842	73.9%
Stillbirths	42 460	0.4%
Elective abortions	1 656 987	17.2%
Inpatient elective abortions	1 390 962	14.4%
Outpatient medical abortions	266 025	2.8%
Therapeutic abortions	93 449	1.0%
<22 weeks after the LMP	69 364	0.7%
≥22 weeks after the LMP	24 085	0.2%
Total abortions[a]	1 830 965	19.0%
Spontaneous abortions	407 925	4.2%
Ectopic pregnancies	108 529	1.1%
Others[b]	131 122	1.4%
Total pregnancy episodes	9 647 843	
Total pregnant women	6 230 200	
Twin pregnancies (live births)	119 404	1.7%

[a]Including "other abortion" type.

[b]Hydatidiform mole or other abnormal products of conception.

equal to the pregnancy start date calculated in the PMSI database for 78.7% of pregnancies and did not differ by more than 1 gestational week for 97.3% of pregnancies.

When taking into account multiple births and pregnancies associated with unknown mother identifiers, the algorithm missed only 0.05% of all live births declared in 2014 (supplementary table 3). The number of pregnancies ending in therapeutic abortions was higher than the official number of authorized abortions. The proportion of unknown mother identifiers was the highest for elective abortions (8% in 2014).

Linkage between maternal and neonatal data was available only for public hospitals in 2011 and has been available for both public and private hospitals since 2012. Linkage rates increased between 2012 and 2015 from 88.5% to 95.2% (Table 3). Linkage rates were 5 points higher for live births than for stillbirths or therapeutic abortions after 22 weeks during the period 2011 to 2015.

3.2 | Antiepileptic drug use

Over the study period, 7 559 701 pregnancies, representing 4 900 139 pregnant women, met the inclusion and exclusion criteria. In this population, 6.7 per 1000 pregnancies were exposed to AEDs: 3.2 to older AEDs and 4.0 to newer AEDs (supplementary table 4). The most commonly used older AEDs were clonazepam, valproic acid, carbamazepine, and phenobarbital with prevalence rates of 1.5, 1.1, 0.6, and 0.1, respectively. The most commonly used newer AEDs were lamotrigine, pregabalin, levetiracetam, topiramate, gabapentin, and oxcarbazepine with prevalence rates of 1.9, 1.0, 0.6, 0.4, 0.3, and 0.2, respectively. Prevalence rates were < 0.1 for the remaining AEDs. Among pregnancies ending in a live birth, 6.3 per 1000 pregnancies were exposed to AEDs.

Exposure to valproic acid, carbamazepine, clonazepam, and pregabalin decreased after the first trimester of pregnancy, while exposure to lamotrigine and levetiracetam remained stable throughout pregnancy (Figure 1).

The number of pregnancies exposed to older AEDs decreased over the study period (−69.4%) (Figure 2), mainly driven by the declining use of clonazepam, valproic acid, and phenobarbital (Figure 3). The use of newer AEDs increased (+73.4%) concomitantly with this decreased use of older AEDs, and newer AEDs became more commonly used than older AEDs after 2010. In particular, the use of levetiracetam, pregabalin, and lamotrigine rapidly increased over the study period. The proportion of women with epilepsy as an LTD among newer AED users decreased from 29.7% to 23.3% between 2007 and 2014 (Figure 4).

Using the 5th or 95th percentile of gestational age instead of the median did not dramatically change the results (Figure 2).

TABLE 3 Linkage rates between maternal and neonatal data for all births (live births, stillbirths, and therapeutic abortions ≥22 weeks after the LMP) by calendar year

	Live Birth	Stillbirth	Therapeutic Abortion ≥ 22 Weeks after the LMP	Total Births
2011	57.9%	55.7%	60.6%	57.9%
2012	88.5%	81.3%	78.9%	88.5%
2013	91.8%	86.4%	85.0%	91.8%
2014	93.9%	88.5%	88.6%	93.8%
2015	95.2%	89.7%	91.0%	95.2%
Total 2011–2015	85.3%	79.9%	80.5%	85.3%

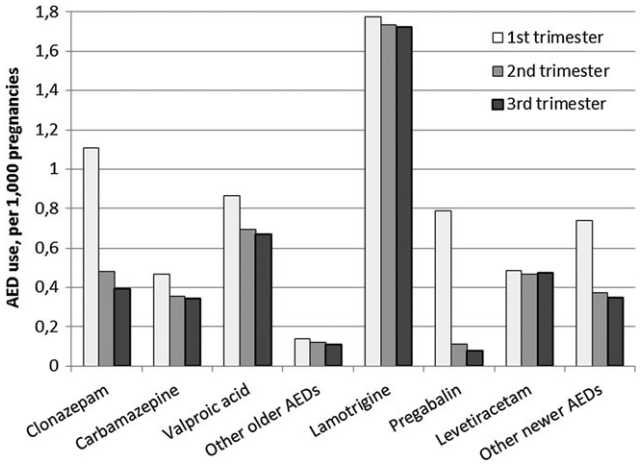

FIGURE 1 Proportion of pregnancies exposed to the most commonly used AEDs according to trimester of pregnancy

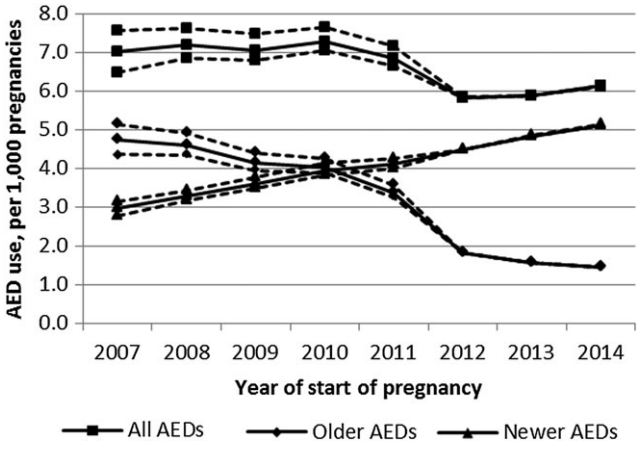

FIGURE 2 Proportion of pregnancies exposed to all types of AEDs and to older and newer AEDs.
Dotted lines represent the proportion of pregnancies exposed to AED using the 5th and 95th percentile of gestational age instead of the median (sensitivity analysis)

4 | DISCUSSION

4.1 | Algorithm

This is the first algorithm to be developed in order to identify pregnancies from the French health care databases. As the French health care databases cover almost all of the French population, this algorithm is a useful tool to conduct studies concerning rare drug exposures or maternal conditions and rare maternal or neonatal outcomes. The algorithm captured all types of pregnancy outcome (live birth, stillbirth, elective abortion, therapeutic abortion, spontaneous abortion, and ectopic pregnancy), allowing not only live births but also other pregnancy outcomes to be included in such studies.

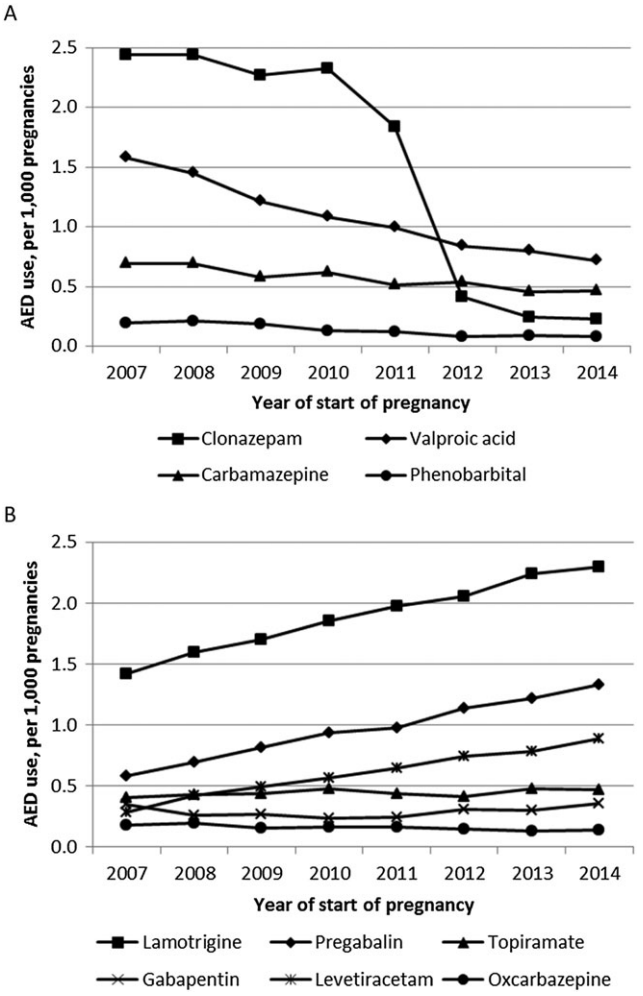

FIGURE 3 Proportion of pregnancies exposed to the most commonly used older (A) and newer (B) AEDs

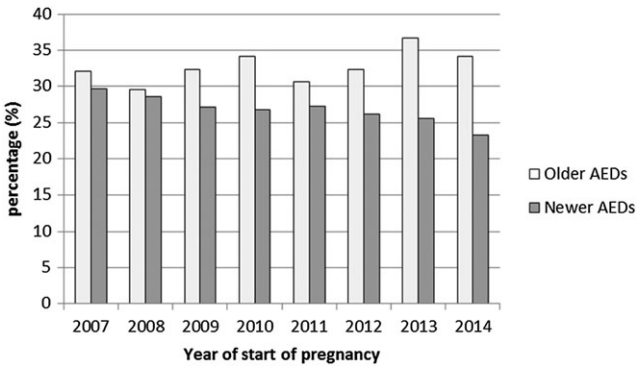

FIGURE 4 Proportion of women with epilepsy LTD status among newer and older AED users after excluding clonazepam from the analysis.
Clonazepam was excluded from the analysis because off-label use was common until the French health authorities took measures to limit off-label use in November 2011

The number of pregnancies identified by the algorithm was compared with official data for live births and therapeutic abortions, demonstrating a high degree of agreement for live births, with a difference of only 2000 pregnancies in 2014. This difference can probably be explained by home deliveries, although codes related to examination after delivery outside hospital were included in the algorithm. Such a comparison was not possible for elective abortions.

The algorithm accurately identified the time period during which a woman was pregnant, as shown by the high concordance rate between the estimated conception date and the pregnancy start date calculated by the algorithm, as a result of the availability of exact pregnancy outcome dates since 2009 and gestational age since March 2010 for births and March 2012 for other outcomes. Unlike many other claims databases, gestational age is directly available in the PMSI database, without the need for linkage to other administrative data, such as vital records.[26] Accurate identification of pregnancy episodes from March 2010 for births and March 2012 for other outcomes should limit misclassification of medication exposure during pregnancy, especially during critical trimesters or months of pregnancy.

Linkage between maternal and neonatal data has been available in the PMSI since 2011 for births ≥22 weeks after the LMP, including stillbirths and therapeutic abortions. The linkage rate was greater than 95% in 2015. Deterministic linkage was possible without the need for probabilistic linkage. This linkage is essential to study the effects of medication exposure or a given condition during pregnancy on neonatal outcomes. The only nationwide study using this linkage published to date was designed to assess the association between maternal gestational diabetes and the risk of adverse neonatal outcomes, such as perinatal death, asphyxia, macrosomia, etc.[15]

This study presents 2 main limitations. First of all, although health care claims data can be exhaustive, readily available, and reasonably inexpensive, making them attractive for large-scale studies, they are not designed for research purposes, unlike registries or, more generally, ad hoc prospectively collected databases, and can be susceptible to misclassification.[20,27] In our study, the proposed algorithm was based on ICD-10 diagnosis codes and medical procedure codes that may be subject to coding errors, and data from medical records could not be used to validate pregnancy outcomes. For instance, the excess number of therapeutic abortions identified by the algorithm could be explained by the omission of the diagnosis code "Problems related to unwanted pregnancy". However, as the PMSI database is used for planning and funding purposes and is subject to coding quality control, coding errors should therefore be limited.

In addition, some pregnancies may not have been identified, particularly anonymized abortions, which can be requested by minors[28] and which represented up to 8% of all elective abortions in 2014. Spontaneous abortions which are not managed in hospital were also missed: the proportion of spontaneous abortions identified with the algorithm was 4.2%, while spontaneous abortion occurs in approximately 15% of all clinically recognized pregnancies.[29] Finally, stillbirths might have been slightly overestimated before March 2011 because therapeutic abortions ≥22 weeks after the LMP for maternal indications, which are far less common than therapeutic abortions for fetal indications, cannot be distinguished from stillbirths.

A second limitation of this study is that exact pregnancy outcome dates were not available before 2009, which could result in imprecise pregnancy start dates. However, the overall prevalence of AED use in our study did not differ by more than 1.3% when the first day or last

day of the month of discharge was used instead of the 15th day. Another source of uncertainty is the absence of recording of gestational age before 2010 for births and 2013 for other outcomes, requiring for instance the use of median gestational ages observed after 2013 in the PMSI database: exposure misclassification could not be ruled out, especially for drugs that are not used chronically like AEDs.[30] In particular, preterm deliveries could not be identified when gestational age was missing. Assigning the same median gestational age to all preterm or full-term live births therefore resulted in too long durations of pregnancy for preterm deliveries. However, the 5th percentile of gestational age was used in a sensitivity analysis and did not substantially change the results.

4.2 | Antiepileptic drug use

This algorithm was implemented in a population of almost 5 million women starting a pregnancy between 2007 and 2014. Over the study period, 6.7 per 1000 pregnancies were exposed to an AED, compared with 8.3 in a previous study based on a small sample of the SNIIRAM database.[16] The prevalence rate, restricted to live births, was 6.3 per 1000 pregnancies, which was higher than those observed in 7 European regions, with prevalence rates ranging from 4.3 in The Netherlands to 6.0 in Wales.[31] However, these prevalence rates cannot be compared directly, as the indications for AEDs vary from 1 country to another. An American study based on administrative health plan data found that 2% of women who gave birth between 2001 and 2007 were exposed to an AED during pregnancy, but mainly for the treatment of psychiatric or pain disorders.[32]

The decreased use of older AEDs and the increased use of newer AEDs between 2007 and 2014 were in line with worldwide trends.[32-34] In particular, the proportion of pregnancies exposed to valproic acid, the most teratogenic AED,[18] decreased between 2007 and 2014.[19,33-36] These trends could be explained by changes in practice guidelines and improved medical knowledge, but other explanations such as changes in population cannot be ruled out. However, valproic acid use during pregnancy remained high in France, particularly during the first trimester of pregnancy, corresponding to the period of greatest risk for the teratogenic effects of medications. As the indication for which a drug is prescribed is not available in the DCIR database, we assessed AED use among women with epilepsy LTD status and observed a decrease in the proportion of women with epilepsy among newer AEDs users between 2007 and 2014, suggesting that the increased use of newer AEDs could be partially explained by a growing use of these drugs in indications other than epileptic disorders, as already documented in Italy and Denmark.[37,38]

5 | CONCLUSION

We have developed an algorithm based on claims data with a number of key strengths for the study of medication use and safety in pregnancy research, especially for pregnancies ending more than 22 weeks after the LMP: the availability of a large study population, accurate calculation of pregnancy outcome and start dates since March 2010, and availability of linkage between maternal and neonatal data since January 2011.

ETHICS STATEMENT

The authors state that no ethical approval was needed.

ACKNOWLEDGEMENTS

The authors thank A. Saul and G. Maura for assistance in writing the manuscript, and P. Berveiller, A. Serfaty and M. Tournaire for providing us with their comments and suggestions on the pregnancy algorithm.

CONFLICT OF INTEREST

The authors are employees of the French National Health Insurance (CNAMTS), the French National Agency for Medicines and Health Products Safety (ANSM), or belong to the French National Institute of Health and Medical Research (INSERM), and have no conflicts of interest with the Pharmaceutical Industry.

ORCID

Pierre-Olivier Blotière http://orcid.org/0000-0002-6857-6781
Alain Weill http://orcid.org/0000-0001-8687-9092

REFERENCES

1. Manson JM, McFarland B, Weiss S. Use of an automated database to evaluate markers for early detection of pregnancy. *Am J Epidemiol*. 2001;154(2):180-187.
2. Devine S, West S, Andrews E, et al. The identification of pregnancies within the general practice research database. *Pharmacoepidemiol Drug Saf*. 2010;19(1):45-50. https://doi.org/10.1002/pds.1862
3. Hardy JR, Holford TR, Hall GC, Bracken MB. Strategies for identifying pregnancies in the automated medical records of the general practice research database. *Pharmacoepidemiol Drug Saf*. 2004;13(11):749-759. https://doi.org/10.1002/pds.935
4. Hornbrook MC, Whitlock EP, Berg CJ, et al. Development of an algorithm to identify pregnancy episodes in an integrated health care delivery system. *Health Serv Res*. 2007;42(2):908-927. https://doi.org/10.1111/j.1475-6773.2006.00635.x
5. Kuklina EV, Whiteman MK, Hillis SD, et al. An enhanced method for identifying obstetric deliveries: implications for estimating maternal morbidity. *Matern Child Health J*. 2008;12(4):469-477. https://doi.org/10.1007/s10995-007-0256-6
6. Palmsten K, Huybrechts KF, Mogun H, et al. Harnessing the Medicaid Analytic eXtract (MAX) to evaluate medications in pregnancy: design considerations. *PloS One*. 2013;8(6):e67405. https://doi.org/10.1371/journal.pone.0067405
7. Naleway AL, Gold R, Kurosky S, et al. Identifying pregnancy episodes, outcomes, and mother-infant pairs in the Vaccine Safety Datalink. *Vaccine*. 2013;31(27):2898-2903. https://doi.org/10.1016/j.vaccine.2013.03.069.
8. Ailes EC, Simeone RM, Dawson AL, Petersen EE, Gilboa SM. Using insurance claims data to identify and estimate critical periods in pregnancy: an application to antidepressants. *Birt Defects Res a Clin Mol Teratol*. 2016;106(11):927-934. https://doi.org/10.1002/bdra.23573
9. Fernandez H, Legendre G, Blein C, Lamarsalle L, Panel P. Tubal sterilization: pregnancy rates after hysteroscopic versus laparoscopic sterilization in France, 2006-2010. *Eur J Obstet Gynecol Reprod Biol*. 2014;180:133-137. https://doi.org/10.1016/j.ejogrb.2014.04.043
10. Le Meur N, Gao F, Bayat S. Mining care trajectories using health administrative information systems: the use of state sequence analysis

to assess disparities in prenatal care consumption. *BMC Health Serv Res*. 2015;15(1):200. https://doi.org/10.1186/s12913-015-0857-5

11. Chantry AA, Deneux-Tharaux C, Bonnet M-P, Bouvier-Colle M-H, Pregnancy-related ICU. Admissions in France: trends in rate and severity, 2006-2009. *Crit Care Med*. 2015;43(1):78-86. https://doi.org/10.1097/CCM.0000000000000601.
12. Raguideau F, Mezzarobba M, Zureik M, Weill A, Ricordeau P, Alla F. Compliance with pregnancy prevention plan recommendations in 8672 French women of childbearing potential exposed to acitretin. *Pharmacoepidemiol Drug Saf*. 2015;24(5):526-533. https://doi.org/10.1002/pds.3763
13. Goueslard K, Cottenet J, Mariet A-S, et al. Early cardiovascular events in women with a history of gestational diabetes mellitus. *Cardiovasc Diabetol*. 2016;15(1):15. https://doi.org/10.1186/s12933-016-0338-0
14. Weill A, Dalichampt M, Raguideau F, et al. Low dose oestrogen combined oral contraception and risk of pulmonary embolism, stroke, and myocardial infarction in five million French women: cohort study. *BMJ*. 2016;353:i2002.
15. Billionnet C, Mitanchez D, Weill A, et al. Gestational diabetes and adverse perinatal outcomes from 716,152 births in France in 2012. *Diabetologia*. 2017;60(4):636-644. https://doi.org/10.1007/s00125-017-4206-6
16. Demailly R, Escolano S, Quantin C, Tubert-Bitter P, Ahmed I. Prescription drug use during pregnancy in France: a study from the national health insurance permanent sample. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1126-1134. https://doi.org/10.1002/pds.4265
17. Bezin J, Duong M, Lassalle R, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2017;26(8):954-962. https://doi.org/10.1002/pds.4233
18. Hernández-Díaz S, Smith CR, Shen A, et al. Comparative safety of antiepileptic drugs during pregnancy. *Neurology*. 2012;78(21):1692-1699. https://doi.org/10.1212/WNL.0b013e3182574f39
19. Christensen J, Grønborg TK, Sørensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *Jama*. 2013;309(16):1696-1703. https://doi.org/10.1001/jama.2013.2270
20. Tuppin P, Rudant J, Constantinou P, et al. Value of a national administrative database to guide public decisions: from the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique*. 2017;65(Suppl 4):S149-S167. https://doi.org/10.1016/j.respe.201705.004
21. Jackson E, Glasier A. Return of ovulation and menses in postpartum nonlactating women: a systematic review. *Obstet Gynecol*. 2011;117(3):657-662. https://doi.org/10.1097/AOG.0b013e31820ce18c
22. Ministère des affaires sociales et de la santé. Manuel des groupes homogènes de malades, version 2016 de la classification. 2016. Available at: http://solidarites-sante.gouv.fr/fichiers/bos/2016/sts_20160005_0001_p000.pdf. Accessed November 23, 2017.
23. Quantin C, Cottenet J, Vuagnat A, et al. Quality of perinatal statistics from hospital discharge data: comparison with civil registration and the 2010 National Perinatal Survey. *J Gynecol Obstet Biol Reprod (Paris)*. 2014;43(9):680-690. https://doi.org/10.1016/j.jgyn.2013.09.004
24. Pierron A, Revert M, Goueslard K, et al. Evaluation of the metrological quality of the medico-administrative data for perinatal indicators: a pilot study in 3 university hospitals. *Rev Epidemiol Sante Publique*. 2015;63(4):237-246. https://doi.org/10.1016/j.respe.2015.05.001
25. Beghi E. Efficacy and tolerability of the new antiepileptic drugs: comparison of two recent guidelines. *Lancet Neurol*. 2004;3(10):618-621. https://doi.org/10.1016/S1474-4422(04)00882-8
26. Margulis AV, Palmsten K, Andrade SE, et al. Beginning and duration of pregnancy in automated health care databases: review of estimation methods and validation results. *Pharmacoepidemiol Drug Saf*. 2015;24(4):335-342. https://doi.org/10.1002/pds.3743
27. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Curr Epidemiol Rep*. 2014;1(4):175-185. https://doi.org/10.1007/s40471-014-0027-z
28. Legifrance. Décret n°2002–799 du 3 mai 2002 relatif à la prise en charge anonyme et gratuite des interruptions volontaires de grossesse pratiquées sur des mineures sans consentement parental. Available at: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000226535. Accessed November 23, 2017.
29. American College of Obstetricians and Gynecologists. ACOG practice bulletin. Management of recurrent pregnancy loss. Number 24, February 2001. (Replaces Technical Bulletin Number 212, September 1995). American College of Obstetricians and Gynecologists. *Int J Gynaecol Obstet*. 2002;78:179-190.
30. Toh S, Mitchell AA, Werler MM, Hernández-Díaz S. Sensitivity and specificity of computerized algorithms to classify gestational periods in the absence of information on date of conception. *Am J Epidemiol*. 2008;167(6):633-640. https://doi.org/10.1093/aje/kwm367.
31. Charlton R, Garne E, Wang H, et al. Antiepileptic drug prescribing before, during and after pregnancy: a study in seven European regions. *Pharmacoepidemiol Drug Saf*. 2015;24(11):1144-1154. https://doi.org/10.1002/pds.3847
32. Bobo WV, Davis RL, Toh S, et al. Trends in the use of antiepileptic drugs among pregnant women in the US, 2001-2007: a medication exposure in pregnancy risk evaluation program study. *Paediatr Perinat Epidemiol*. 2012;26(6):578-588. https://doi.org/10.1111/ppe.12004
33. Wen X, Meador KJ, Hartzema A. Antiepileptic drug use by pregnant women enrolled in Florida Medicaid. *Neurology*. 2015;84(9):944-950. https://doi.org/10.1212/WNL.0000000000001304
34. Eurap Study Group. Utilization of antiepileptic drugs during pregnancy: comparative patterns in 38 countries based on data from the EURAP registry. *Epilepsia*. 2009;50(10):2305-2309. https://doi.org/10.1111/j.1528-1167.2009.02093.x
35. Vajda FJE, Hollingworth S, Graham J, et al. Changing patterns of antiepileptic drug use in pregnant Australian women. *Acta Neurol Scand*. 2010;121(2):89-93. https://doi.org/10.1111/j.1600-0404.2009.01260.x
36. Man S-L, Petersen I, Thompson M, Nazareth I. Antiepileptic drugs during pregnancy in primary care: a UK population based study. *PloS One*. 2012;7(12):e52339. https://doi.org/10.1371/journal.pone.0052339
37. Savica R, Beghi E, Mazzaglia G, et al. Prescribing patterns of antiepileptic drugs in Italy: a nationwide population-based study in the years 2000-2005. *Eur J Neurol*. 2007;14(12):1317-1321. https://doi.org/10.1111/j.1468-1331.2007.01970.x
38. Tsiropoulos I, Gichangi A, Andersen M, Bjerrum L, Gaist D, Hallas J. Trends in utilization of antiepileptic drugs in Denmark. *Acta Neurol Scand*. 2006;113(6):405-411. https://doi.org/10.1111/j.1600-0404.2006.00639.x

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Blotière P-O, Weill A, Dalichampt M, et al. Development of an algorithm to identify pregnancy episodes and related outcomes in health care claims databases: An application to antiepileptic drug use in 4.9 million pregnant women in France. *Pharmacoepidemiol Drug Saf*. 2018;27:763–770. https://doi.org/10.1002/pds.4556