

Trajectoires individuelles d'exposition aux psychotropes au cours de la grossesse et partitionnement en fonction du profil d'exposition : utilisation des K-means pour données longitudinales

Caroline Hurault-Delarue¹, Cécile Chouquet², Nicolas Savy², Isabelle Lacroix¹, Christine Damase-Michel¹

¹ Service de Pharmacologie médicale, CHU de Toulouse, Université de Toulouse, INSERM 1027
Faculté de Médecine 37, allées Jules Guesde 31000 Toulouse

caroline.delarue-hurault@univ-tlse3.fr

isabelle.lacroix@univ-tlse3.fr

christine.damase-michel@univ-tlse3.fr

² Université Paul Sabatier, Institut de Mathématiques de Toulouse, UMR 521
118, route de Narbonne F-31062 Toulouse Cedex 9

cecile.chouquet@math.univ-toulouse.fr

nicolas.savy@math.univ-toulouse.fr

Résumé. La dimension longitudinale de l'exposition médicamenteuse n'est que rarement considérée dans les études de pharmaco-épidémiologie. Sa prise en compte est pourtant importante dans certaines situations comme la grossesse. L'utilisation des bases de données administratives de prescriptions de médicaments permet la prise en compte de cet aspect.

L'objectif de cette étude est de décrire les trajectoires individuelles d'exposition des femmes aux psychotropes au cours de leur grossesse et de constituer des groupes de patientes aux profils d'exposition proches.

Les données de prescriptions et délivrances de psychotropes de la cohorte EFEMERIS ont été utilisées, concernant plus de 3700 femmes enceintes. La nature hétérogène de ces données de prescriptions (en terme de dosage, de principes actifs, ...) nous a amené à calculer le nombre de Doses Définies Journalières (DDD) par femme et par mois de grossesse. Cette mesure standardisée reflète l'exposition des femmes aux psychotropes au cours de la grossesse, représentée sous forme de trajectoire individuelle. La méthode de classification de type K-means adaptée aux données longitudinales (KML) a été appliquée sur ces trajectoires dans le but de constituer des groupes de femmes homogènes en termes d'exposition (clusterisation).

Cette approche a permis de décrire plus précisément l'exposition des femmes enceintes aux psychotropes et de constituer des « profils » d'exposition. Les résultats de cette première étude nous encourageant à utiliser cette méthode pour d'autres études de pharmaco-épidémiologie. Par la suite, la relation entre « profils » d'exposition aux psychotropes in utero et anomalies de développement psychomoteur de l'enfant sera étudiée.

Mots-clés : Données longitudinales, KML, médicaments, grossesse

Abstract.

The longitudinal dimension of drug exposure is rarely considered in studies of pharmaco-epidemiology. However, taking into account this dimension is important in some situations such as pregnancy. The use of administrative databases of prescriptions drugs allows taking into account this parameter.

The aim of the present study was to describe individual trajectories of exposure to psychotropic drugs during pregnancy and to identify homogeneous groups of women exposed.

Data on psychotropic drug prescriptions in EFEMERIS cohort was used: it concerns more than 3,700 pregnant women. The heterogeneity of these data (due to the various drug dosages and posologies) led us to calculate the number of Defined Daily Dose (DDD) by woman and by month of pregnancy. For each woman, all these standardized measures reflect the exposure to psychotropic drugs all along pregnancy and can be represented as individual trajectories. K-means method specially designed to longitudinal data (KML) was applied on these trajectories in order to classify pregnant women in homogeneous exposure groups (clusterization).

This new mathematical approach elicits a more precise description of psychotropic exposed women all along their pregnancy and allows us to constitute psychotropic exposure profiles. The results of this first study encourage us to use this method for further studies of pharmacoepidemiology. Thereafter, the relationship between exposure to psychotropic drugs in utero and potential abnormalities of psychomotor development of the child will be studied.

Keywords : Longitudinal data, KML method, drugs, pregnancy

CONTEXTE

La dimension longitudinale de l'exposition médicamenteuse n'est que rarement considérée dans les études de pharmaco-épidémiologie. Sa prise en compte est pourtant importante dans certaines situations comme par exemple la grossesse. En effet, les risques sur l'enfant (malformatifs, foeto-toxiques, retard du développement psychomoteur) liés à une exposition médicamenteuse au cours de la grossesse varient en fonction de la dose mais également du temps, de la période et de la durée d'exposition.

L'utilisation des bases de données administratives, recueillant de façon prospective les données de prescriptions au cours du temps, permet la prise en compte de cet aspect et la reconstitution des trajectoires individuelles d'exposition médicamenteuse.

Les psychotropes représentent une classe de médicaments prescrite assez largement aux femmes au cours de leur grossesse (près de 7%). Les schémas posologiques (dose, durée, associations) de ces médicaments sont variés.

L'objectif de cette étude est de décrire les trajectoires individuelles d'exposition des femmes aux psychotropes au cours de leur grossesse et de constituer des groupes de patientes (« clusters ») aux profils d'exposition proches.

METHODES

Les données de prescriptions et délivrances de médicaments des femmes au cours de la grossesse (issues des bases de données administratives de prescriptions de l'Assurance Maladie) de la base de données EFEMERIS (Evaluation chez la Femme Enceinte des Médicaments et de leurs RISques, présentée par Lacroix (2009)), mise en place en Haute-Garonne en 2004, ont été utilisées pour cette étude.

Cette étude concerne les femmes ayant accouché du 1^o juillet 2004 au 31 décembre 2010, incluses dans cette cohorte, ayant reçu au cours de leur grossesse au moins une prescription de psychotropes.

Les différentes classes de psychotropes prises en compte sont les suivantes (selon la classification ATC) :

- Antipsychotiques/neuroleptiques (N05A),
- Anxiolytiques/hypnotiques/benzodiazépines (N05B, N05C),
- Antidépresseurs (N06A, N06C),
- Antiépileptiques (N03).

EFEMERIS contient les noms exacts des spécialités remboursables prescrites et délivrées aux femmes au cours de leur grossesse ainsi que leur dosage et les dates de chaque délivrance.

La nature de ces données de prescriptions nous a amené à faire des choix méthodologiques que l'on peut résumer en trois points :

1. Les psychotropes peuvent être de natures très différentes (grand nombre de spécialités). La base de données ne mentionne pas la posologie (doses et modalités d'administration des médicaments) et les durées exactes de traitement définies par le prescripteur ; or celles-ci peuvent varier considérablement d'un psychotrope à l'autre.

Pour pallier cette hétérogénéité des données de prescription, la **DDD** (Defined Daily Dose) a donc été utilisée pour créer une variable de mesure d'exposition aux psychotropes. La DDD, fixée par le WHO Collaborating Centre (2013), correspond à la posologie journalière théorique d'une substance

médicamenteuse utilisée dans son indication principale chez l'adulte. La méthodologie de la DDD permet non seulement de convertir les quantités physiques de médicaments (nombre de comprimés par exemple) en une unité de mesure standard, mais aussi de regrouper les données sur l'utilisation de médicaments ou de classes de médicaments différents.

2. Pour prendre en compte l'aspect longitudinal des données, les mesures d'exposition basées sur la DDD ont été regroupées par femme et considérées comme des **trajectoires individuelles d'exposition**.

3. La distinction de différents profils d'exposition aux psychotropes est nécessaire pour la réalisation d'études sur l'évaluation du risque lié à la prise de psychotropes au cours de la grossesse.

La **constitution de groupes homogènes** de femmes en terme de trajectoires d'exposition aux psychotropes au cours de la grossesse (clusterisation) a donc été réalisée grâce à une méthode de classification non supervisée (ou clustering) de type K-means adaptée aux données longitudinales (KML) développée par Génolini et Falissard (2009, 2011). La méthode des K-means, basée sur le principe de ré-allocation dynamique des individus à des centres de classes, est une méthode itérative ; après initialisation des centres de gravité, l'algorithme répète 2 opérations jusqu'à la convergence d'un critère :

- 1° opération « Espérance » : le centre de gravité de chaque groupe est calculé
- 2° opération « Maximisation » : chaque individu est affecté à la classe dont le centre est le plus proche.

Chaque cluster est caractérisé par son centre de gravité et son inertie. L'inertie d'un cluster mesure la concentration des points du cluster autour du centre de gravité. Plus cette inertie est faible, plus petite est la dispersion des points autour du centre de gravité. Les centres de gravité des clusters forment eux aussi un nuage de points caractérisé par l'inertie inter-cluster. L'inertie inter-cluster mesure "l'éloignement" des centres des clusters entre eux. Plus cette inertie est grande, plus les clusters sont bien séparés. L'objectif de cette classification est donc de minimiser l'inertie intra-cluster et maximiser l'inertie inter-cluster. Le choix du nombre de clusters est basé sur le critère de Calinski et Harabatz (1974).

A la suite de cette clusterisation, la description des caractéristiques des différents groupes de femmes constitués a été réalisée.

EFEMERIS contient depuis Janvier 2012 (date postérieure à notre période d'étude) le nombre de boîtes de médicaments délivrées lors du passage à la pharmacie. Cette donnée pouvant avoir une influence non négligeable sur notre classification des femmes, nous en avons tenu compte en simulant le nombre de boîtes délivrées par prescription, selon une loi multinomiale.

RESULTATS

3708 femmes exposées aux psychotropes au cours de leur grossesse sont concernées par cette étude. Le partitionnement en 4 clusters a été choisi suite à la prise en compte du critère de Calinski et Harabatz et de la plausibilité clinique des trajectoires.

Le cluster A regroupe 72,0% de la population. Il se caractérise par un niveau d'exposition "faible" aux psychotropes constant au cours de la grossesse.

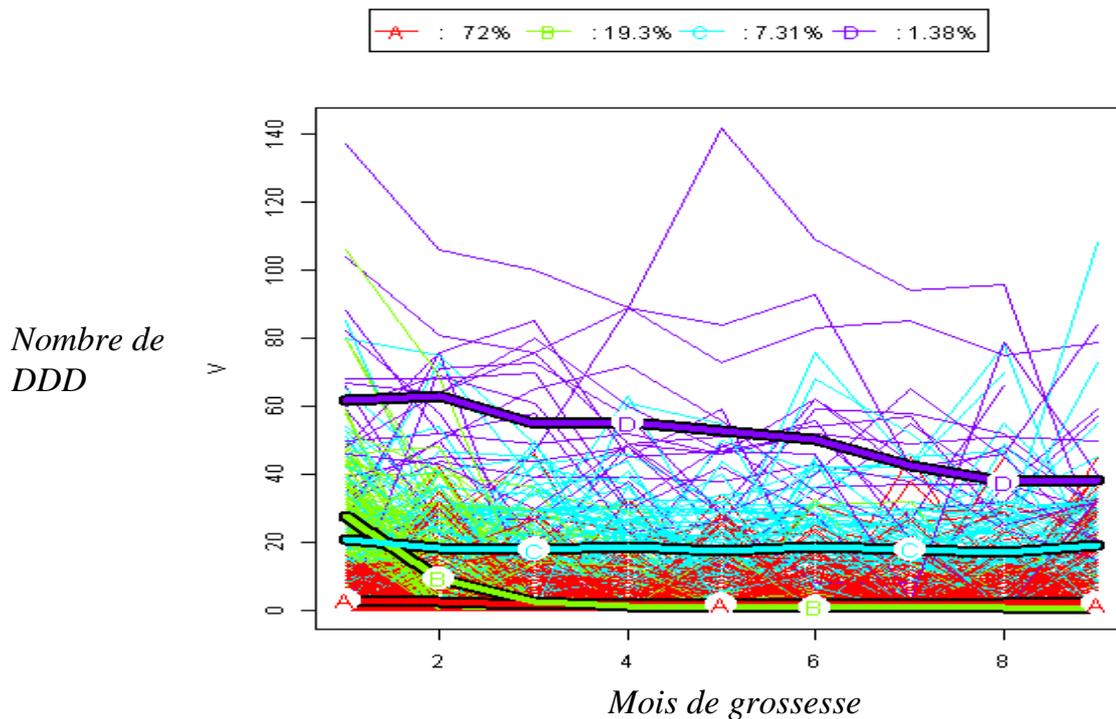
Le cluster B regroupe 19,3% de la population. Il se caractérise par une diminution brutale de l'exposition au cours du premier trimestre de grossesse et une exposition stable et "faible" sur la suite de la grossesse.

Le cluster C regroupe 7,3% de la population. Il se caractérise par un niveau d'exposition "moyen" mais stable tout au long de la grossesse.

Le cluster D regroupe seulement 1,4% de la population. Il se caractérise par un niveau d'exposition

"élevé" mais une tendance à la diminution des doses au cours de la grossesse.

La représentation graphique de ce partitionnement est présentée ci-dessous :



Une analyse descriptive des caractéristiques des femmes par cluster a montré en particulier que les femmes les plus exposées étaient caractérisées par un âge plus élevé et un plus grand nombre d'échographies que les femmes moins exposées.

La clusterisation a également été réalisée sur chacune des classes de psychotropes (antipsychotiques/neuroleptiques, anxiolytiques/hypnotiques/benzodiazépines, antidépresseurs, antiépileptiques).

Une bonne « concordance » entre la clusterisation présentée ci-dessus et celle issue des simulations a été retrouvée ($Kappa = 0,77 [0,75-0,80]$).

CONCLUSION

Cette approche nous a permis de décrire plus précisément l'exposition des femmes enceintes aux psychotropes et de constituer des « profils » d'exposition. Les résultats cohérents de cette première étude nous encourageant à utiliser cette méthode pour d'autres études de pharmaco-épidémiologie.

Dans la continuité de cette première partie, la relation entre « profils » d'exposition aux psychotropes in utero et anomalies de développement psychomoteur fera l'objet d'une nouvelle étude, qui devra prendre en compte plusieurs difficultés méthodologiques telles que la modélisation d'une variable réponse multidimensionnelle et la présence de données manquantes.

BIBLIOGRAPHIE

- [1] Lacroix I, Hurault C, Sarramon MF, Guitard C, Berrebi A, Grau M, et al. (2009), Prescription of drugs during pregnancy: a study using EFEMERIS, the new French database, *Eur J Clin Pharmacol.*, 65(8):839–846.
- [2] WHOCC - Definition and general considerations (2013), Disponible sur: http://www.whooc.no/ddd/definition_and_general_considera/
- [3] Genolini C, Falissard B. (2009), KmL: k-means for longitudinal data, *Comput Stat.*, 25(2):317–328.
- [4] Genolini C, Falissard B. (2011), Kml: A package to cluster longitudinal data, *Comput Stat.*, 104(3):112–121.
- [5] Calinski T, Harabasz J. (1974), A dendrite method for cluster analysis, *Commun Stat - Theory Methods*, 3(1):1–27.