Rapport scientifique final

RENSEIGNEMENTS ADMINISTRATIFS

Identification du projet	
Titre du projet	COmplex Statistical Approaches for modelling and quantifying the impact of drug exposure from medico-administrative data: application to PREGnancy (COSA-PREG)
Coordinateur scientifique du projet (société/organisme - laboratoire ou entité de rattachement)	Pascale TUBERT-BITTER Inserm U1018
Référence de l'appel à projets (nom + année)	AAP Général Prévention 2017
Citez les équipes partenaires (organismes, structures ou laboratoire de rattachement)	
Durée **initiale** du projet	36 mois
Période du projet (date début – date fin)	1er mars 2019 – 28 février 2023
Avez-vous bénéficié d'une prolongation ? Si oui, de quelle durée ?	Oui : 12 mois
Le cas échéant, expliquez succinctement les raisons de cette prolongation.	Le projet comporte 2 axes principalement portés par deux doctorants. Chacun de ces axes a connu du retard, pour une bonne part dû à la situation sanitaire qui a pesé sur leurs conditions de travail, (le doctorant développant le 2ème axe a de plus été très mobilisé par son activité clinique hospitalière pendant toute la crise et les répercutions post-crise). Le travail de la doctorante (axe 1) a également été interrompu par un congé maternité. Le déroulement de la recherche a de plus enrichi le projet avec des éléments qui n'étaient pas initialement prévus, l'identification d'une question inattendue (axe 1), et la mise en œuvre d'une

	stratégie alternative de construction d'un ensemble d'associations (médicaments – effets indésirables) de référence (axe 2).
Identification du rédacteur	
Nom du rédacteur du rapport	Pascale Tubert-Bitter
Fonction au sein du projet *(si différent du Coordinateur scientifique du projet)*	
Téléphone	01 45 59 52 62
Adresse électronique	Pascale.tubert@inserm.fr
Date de rédaction du rapport	27 /04/2023

Indiquer la liste du personnel recruté dans le cadre du projet *(si pertinent)*

Nom	Prénom	Qualifications	Date de recrutement	Durée du contrat (en mois)	Type de contrat (CDD, vacation …)
Manitchoko	Liliane	Biostatistique	23 juin 2019	35 mois	CDD

Donner la liste des autres personnes impliquées dans le projet.

Nom	Prénom	Qualifications	% de son temps consacré au projet
Ahmed	Ismaïl	CR Biostatistique	30
Antoni	Guillemette	IR Biostatistique	10
Escolano	Sylvie	IR Biostatistique	20
Demailly	Romain	Sage femme, épidémiologie	30
Thiébaut	Anne	CR Biostatistique	30
Tubert-Bitter	Pascale	DR Biostatistique	30

Donner la répartition (en pourcentage) des disciplines mobilisées dans le projet.

Il s'agit d'un projet essentiellement méthodologique, avec une large implication de biostatisticiens. Ce projet comporte également une forte composante appliquée, axée sur les effets indésirables des médicaments chez la femme enceinte, avec la mobilisation d'une discipline clinique, la maïeutique. Par ailleurs, la thèse de Liliane Manitchoko (axe 1) était co-encadrée par Jacques Bénichou (CHU Rouen, dans l'équipe, biostatisticien), dont l'article a permis la collaboration avec Michal Abrahamowicz (McGill University) et nous avons bénéficié de la collaboration de Françoise Haramburu (pharmacologue pharmacovigilante, CHU Bordeaux) pour l'axe 2, et de l'expertise de Catherine Quantin (CHU Dijon, dans l'équipe) sur les données de santé.

SYNTHESES ET RAPPORT SCIENTIFIQUE

Synthèse courte

Contexte et objectifs

Les bases médico-administratives sont de plus en plus utilisées en pharmacoépidémiologie. En France, le Système National des Données de Santé (SNDS) couvre la quasi-totalité de la population française et participe à l'essor des études populationnelles à large échelle en apportant des connaissances précieuses pour la prise de décision en santé publique.

La prescription de médicaments pendant la grossesse est courante. Cependant, les informations dont disposent les cliniciens en termes d'effets indésirables sont rares. Les études utilisant le SNDS sont particulièrement intéressantes à cet égard. Ces bases de données contiennent une grande quantité d'informations concernant, notamment, l'exposition aux médicaments remboursés. En revanche, elles se heurtent à des informations limitées sur les facteurs de risque classiques.

L'objectif général de ce projet est d'évaluer des méthodes et modèles statistiques pour réaliser des études pharmaco-épidémiologiques à partir de bases de données médico-administratives. En particulier, il s'articule autour de deux axes méthodologiques : 1) la modélisation de la relation temporelle entre l'exposition médicamenteuse d'intérêt, variable dans le temps, et l'effet consécutif éventuel, et 2) la prise en compte de la confusion mesurée et non mesurée. Pour chacun de ces axes, il s'agit d'essayer d'identifier et d'évaluer les biais induits par les designs et l'analyse statistique (axe 1), et par les méthodes proposées pour corriger les biais de confusion (axe 2). Ce projet comporte également une forte composante appliquée puisqu'il est spécifiquement déployé pour l'axe 2 pour la pharmacoépidémiologie de la grossesse.

Travaux réalisés (méthodologie et résultats)

Pour le 1er axe, l'objectif était d'examiner les performances des schémas d'étude de cohorte et cas-témoins niché (CTN) pour estimer l'association entre une exposition médicamenteuse et le risque d'évènement indésirable, en particulier pour une exposition variable dans le temps.

Les résultats de l'étude de simulation que nous avons réalisée ont montré que les estimations obtenues de l'analyse de la cohorte entière étaient non biaisées dans tous les scénarios considérés. Cependant, les estimations issues des analyses CTN (avec l'approximation d'Efron) étaient substantiellement biaisées, en particulier lorsqu'un seul témoin était apparié à chaque cas. Ce biais augmentait avec la proportion d'évènements.

Pour comprendre d'où venaient les biais dans les résultats des analyses CTN et y remédier éventuellement, nous avons dans un premier temps analysé les données CTN avec la méthode de Firth adaptée à la régression logistique conditionnelle pour la réduction des biais. Nous avons constaté une diminution considérable des biais dans les estimations suggérant ainsi que les biais observés pourraient provenir de données éparses.

Cependant, certains arguments nous ont fait comprendre que cette explication n'était pas satisfaisante. En effet, après avoir augmenté la taille des cohortes simulées (résultats non présentés dans ce mémoire) et surtout quel que soit le nombre d'évènements, les biais persistaient alors qu'en principe

plus la proportion d'évènements augmente plus le risque de données éparses diminue. Nous avons donc, dans un second temps, pris en compte les temps d'évènement exæquos (pouvant apparaître dans les strates) en utilisant différentes méthodes : les approximations de Breslow et d'Efron (cette dernière étant celle utilisée par défaut dans les premières analyses CTN et de la cohorte entière), la méthode exacte et l'approche *ccwc* modifiée. Nous avons constaté que les estimations CTN avec l'approximation de Breslow étaient davantage biaisées qu'avec l'approximation d'Efron, tandis qu'avec la méthode exacte ou l'approche *ccwc* modifiée les estimations étaient presque exemptes de biais et assez proches de celles de l'analyse de la cohorte entière. Il est ainsi apparu que les biais initialement observés résultaient vraisemblablement de la présence de temps d'évènement ex-æquos dans les strates. Ces ex-æquos sont présents dans la cohorte entière mais relativement "dilués" de sorte que leur prise en compte impacte peu l'estimation de l'effet de l'exposition.

Quel que soit le scénario considéré, les estimations basées sur l'analyse de la cohorte entière ont présenté une précision systématiquement plus faible que celles des analyses CTN. Les estimations issues de l'analyse de la cohorte entière avaient un taux de recouvrement satisfaisant tandis que celles issues de l'analyse CTN présentaient de faibles taux de recouvrement, en particulier avec les approximations de Breslow et d'Efron pour des scénarios où des biais considérables avaient été observés. Tous ces résultats ont été observés autant pour une exposition variable que fixe dans le temps.
Il semble que la détérioration potentielle des estimations issues de l'analyse CTN en présence de nombreux évènements ex-æquos soit méconnue des modélisateurs et que ce problème ne soit pas suffisamment pris en compte dans la pratique. En particulier, les chercheurs devraient être conscients des limites possibles des options par défaut utilisées dans les logiciels statistiques disponibles pour mettre en œuvre la régression logistique conditionnelle.

Le 2ème axe était décliné en trois volets étroitement connectés. La méthodologie d'évaluation ici est structurée par l'analyse des données réelles, l'échantillon permanent (EGB) du SNDS pour les deux premiers, et le SNDS complet pour le 3ème volet. Nous avons voulu (i) nous doter d'un outil, la constitution d'un ensemble de référence d'associations médicamenteuses pour la grossesse, à même de permettre l'évaluation des méthodes étudiées, dans un premier temps pour la prématurité (dans une forme de preuve de concept), puis élargi ; (ii) mettre en œuvre les scores synthétiques (score de propension, score pronostique) pour la réduction des biais de confusion, en grande dimension, appliquée à la prématurité ; (iii) mettre en œuvre « grandeur réelle » une analyse par variable instrumentale pour pallier la confusion non mesurée d'une association de référence. Les deux premiers volets ont été abordés en développant une méthodologie de détection de signaux basée sur l'utilisation des scores en grande dimension. Cette méthodologie, déclinée en 21 modalités, est appliquée pour identifier des médicaments (parmi les 400 testés) possiblement associés à une augmentation du risque de prématurité. Un critère de détection dérivé de ces scores, le p-RD (p-value relative decrease), est proposé et sa pertinence est évaluée grâce à l'ensemble de référence créé à partir d'une revue manuelle de la littérature de MEDLINE complétée par une expertise pharmacologique pour les signaux générés par l'une des méthodes. Les résultats montrent que l'utilisation du p-RD dérivé du score pronostique permet de mieux prendre en compte la confusion et de limiter les biais d'indication et protopathique.

Pour son extension, le p-RD est appliqué à vingt pathologies spécifiques ou non de la grossesse et combiné à un requêtage automatisé des mots-clés MeSH des articles de MEDLINE, pour obtenir un ensemble annoté de signaux. Le troisième volet s'appuie sur une association d'intérêt obtenue au cours de cette recherche (1ère volet), l'escitalopram, un antidepresseur, comparé aux autres médicaments de sa classe (inhibiteurs sélectifs de la recapture de la sérotonine, ISRS) et le risque d'accouchement prématuré. La mise en œuvre de l'analyse par variable instrumentale, la préférence du prescripteur, sur toutes les grossesses entre 2012 et 2014 exposées à un ISRS n'a pas permis d'améliorer les estimations obtenues par les analyses classiques.

Perspectives

Les retombées scientifiques de ce projet se situent à deux niveaux. D'une part, en matière de prévention, l'acquisition de connaissances sur les risques médicamenteux de prématurité pendant la grossesse pourra aider les praticiens à prendre des décisions concernant d'éventuels ajustements de traitements. D'autre part, d'un point de vue pharmaco-épidémiologique, la méthodologies étudiées dans le domaine de la grossesse pourrait être transposée à d'autres champs pour évaluer les risques des médicaments, en utilisant le SNDS, quelle que soit la population ou le médicament étudié.

Rapport scientifique

1. Contexte de la recherche

Les études de pharmacoépidémiologie sont de plus en plus souvent réalisées à partir de bases de données médico-administratives volumineuses et complexes. C'est également le cas en France, avec l'utilisation croissante de la base de données du Système National des Données de Santé (SNDS), comme le montrait déjà à l'époque du dépôt du projet une communication de la Caisse Nationale de l'Assurance Maladie (CNAM, Weill et al. 2017). Dans la même lignée, deux plateformes avaient été financées par l'Agence de Sécurité des Médicaments et des produits de santé (ANSM) pour réaliser des études pharmacoépidémiologiques à partir des données du SNDS. L'accès à ces données s'est amplifié depuis avec la création du Health Data Hub, l'accès permanent au SNDS pour les équipes de recherche en santé, et pour la pharmacoépidémiologie, la constitution par la CNAM et l'ANSM du GIS EPI-PHARE.

Ces données contiennent une grande richesse d'informations concernant notamment l'exposition aux médicaments remboursés. Elles permettent également les modélisations statistiques les plus adaptées à rendre compte des variations de l'exposition médicamenteuse et de son effet sur le risque d'un événement de santé dans le temps. En revanche, ces données contiennent peu d'informations sur les facteurs de risque classiques (par exemple, le statut tabagique, la consommation, l'IMC...). Ce manque d'information sur des facteurs de confusion importants peut entraîner une estimation biaisée des risques induits par les médicaments, alors que l'utilisation de ces données contribue de plus en plus aux questions de santé publique. Notre équipe avait été l'une des premières à obtenir l'autorisation d'extraire

les données du SNDS (SNIIRAM alors) pour une recherche méthodologique sur les risques liés aux vaccins chez les jeunes enfants dont l'objectif portait sur la méthode des séries de cas autocontrôlées (Farrington Biometrics 1995), une méthode qui présente l'avantage essentiel d'ajuster implicitement les facteurs de confusion fixes. Cela a lancé notre ligne de recherche sur l'évaluation de la faisabilité et des limites de l'utilisation des données du SNDS en pharmacoépidémiologie (Hanf et al vaccine 2013).

La prescription de médicaments pendant la grossesse est courante. Cependant, les informations dont disposent les prescripteurs en termes d'effets indésirables sont rares, car les essais cliniques impliquent rarement des femmes enceintes. Dans ce contexte, les études de population sont devenues une ressource cruciale. Le projet considère les femmes enceintes pour sa population d'intérêt pour lesquelles notre équipe développe un vaste programme de recherche. En particulier, P. Tubert-Bitter co-pilotait avec C. Quantin des projets financés par la FRM et l'ANR dont l'objectif principal était d'étudier la qualité métrologique des données hospitalières PMSI pour les indicateurs de santé périnatale et d'évaluer certaines méthodologies statistiques pour évaluer les associations entre des médicaments spécifiques et des troubles néonataux en utilisant les données du SNDS (projet ETOMEG, Goueslard K *et al* 2020, Quantin C *et al* 2021, Koladjo *et al* 2018). De plus, I. Ahmed et P. Tubert-Bitter co-encadraient la thèse de doctorat de R. Demailly, sage-femme en maternité lilloise, sur la "Détection des signaux de sécurité des médicaments pendant la grossesse à partir des bases de données médico-administratives" qui implique des approches statistiques de grande dimension. En utilisant l'échantillon de bénéficiaires 1/97e du SNIIRAM, il avait pu décrire l'exposition aux médicaments de prescription pendant la grossesse en France au cours des dernières années (Demailly et al. 2017). La présente recherche s'appuyait sur ces deux projets.

2. Rappels des objectifs

Ce projet comporte une forte composante méthodologique dont l'objectif global est de caractériser, évaluer, développer, calibrer et mettre en œuvre des outils statistiques pour améliorer les études pharmaco-épidémiologiques réalisées à partir de bases de données médico-administratives. Il s'articule autour de deux axes de modélisation et d'évaluation des biais et des stratégies de leur réduction :

(1) la relation temporelle entre l'exposition au médicament et l'événement d'intérêt

(2) la confusion mesurée et non mesurée.

Ce projet méthodologique vise à fournir des approches valables pour l'étude des questions pharmaco-épidémiologiques pendant la grossesse. Plus particulièrement, il se concentre sur l'effet de l'exposition aux médicaments sur la prématurité.

Ce projet s'inscrivait dans l'axe 2.1 de l'appel " Développer la recherche cognitive nécessaire en amont du développement des actions de prévention ". Plus précisément, les questions méthodologiques que nous proposions d'aborder visent à mieux quantifier la relation entre l'exposition aux médicaments et un événement d'intérêt donné à partir de grandes bases de données médico-administratives. La proposition se caractérise par un intérêt méthodologique marqué et une focalisation sur les méthodes

statistiques récemment développées dans le domaine de la pharmaco-épidémiologie. Les recherches sur la prise en compte de la confusion mesurée s'appuient sur l'ensemble des travaux menés par Schneeweiss *et al*. sur le score de propension en grande dimension (Schneeweiss *et al*. 2009). Enfin, la source d'inspiration concernant la confusion non mesurée peut être trouvée dans la littérature relative à l'utilisation de variables instrumentales (Rassen *et al.* 2009). Ce projet visait également l'étude d'associations médicamenteuses spécifiques.

3. Méthodologies utilisées

Les aspects méthodologiques qui sont étudiés peuvent être divisés en deux catégories. La première concerne l'amélioration de la modélisation des effets des médicaments dans le contexte de l'analyse de grandes bases de données médico-administratives telles que le SNDS. Pour mieux comprendre la nature dynamique de l'exposition aux médicaments et sa relation avec les résultats de santé, il est courant en pharmaco-épidémiologie de concevoir des études de cohortes et d'utiliser des techniques d'analyse de survie. La plupart des études pharmaco-épidémiologiques publiées, réalisées à partir de grandes bases de données de demandes de remboursement, ont modélisé cette relation à l'aide de modèles de Cox, avec une exposition aux médicaments dépendante du temps. Néanmoins tous les biais induits par l'analyse ou le design n'ont pas été étudiés de manière approfondie.

Le deuxième aspect méthodologique que nous proposons d'étudier concerne la réduction des biais dus aux variables de confusion mesurées et non mesurées. Malgré la richesse des données présentes dans le SNIIRAM, ce dernier manque d'informations cruciales concernant les facteurs de risque classiques. D'autre part, cette base de données contient des milliers de variables liées aux diagnostics, aux procédures et aux médicaments. Certains auteurs ont émis l'hypothèse qu'en explorant ces variables, il serait possible d'identifier un ensemble de variables indirectes susceptibles de décrire l'état de santé des patients. C'est la raison d'être du score de propension en grande dimension (hdPS) proposé par Schneeweiss *et al*. (Schneeweiss *et al.* 2009). La méthodologie du score de propension a été développée à l'origine pour les études d'observation afin d'imiter les essais cliniques randomisés dans lesquels les patients traités et non traités sont rendus comparables sur toutes les caractéristiques, à l'exception de leur traitement, grâce à une répartition aléatoire dans les groupes de traitement. La randomisation permet donc une interprétation causale de toute association médicament-effet qui pourrait être trouvée dans l'analyse. Dans les études d'observation, les facteurs de confusion, en particulier, empêchent de tirer des conclusions causales. Le score de propension est obtenu comme la probabilité prédite qu'un individu aurait été traité sur la base des caractéristiques observées de cet individu avant le traitement. L'appariement, l'ajustement ou la pondération de l'analyse principale des médicaments et de leurs effets sur la base du score de propension permet d'éliminer (du moins en grande partie) les facteurs de confusion et de mieux s'approcher de l'effet causal du traitement. En bref, l'approche hdPS utilise des évaluations empiriques de la prévalence des variables et de leur association avec l'exposition et les effets pour sélectionner des milliers de variables à inclure dans le score de

propension. Ces variables sont ensuite classées en fonction de leur potentiel de confusion et les variables les mieux classées sont utilisées pour calculer un score de propension.

Malgré son intérêt, le hdPS présente certaines limites. Tout d'abord, le choix du nombre de variables à sélectionner dans le score de propension est laissé à l'utilisateur. Or, ce paramètre a un impact significatif sur les performances des méthodes comme l'a montré une récente étude de simulation basée sur deux sources de données différentes (Franklin *et al*. 2015). Il est donc intéressant de travailler sur des stratégies permettant de le choisir d'une manière plus statistique, soit en optimisant certains critères basés sur la prédiction, soit en optimisant des critères basés sur la sélection de modèles. Un autre inconvénient de la méthode hdPS est que la sélection des facteurs de confusion potentiels est effectuée de manière univariée : chaque covariable est étudiée séparément des autres. Or, la relation entre les facteurs de confusion potentiels peut être très complexe. Il est donc probablement avantageux d'utiliser des algorithmes multivariés adaptés à des contextes de grande dimension (Lee *et al*. 2010). En particulier, les régressions basées sur le lasso ou les méthodes directement empruntées à l'apprentissage automatique telles que les arbres boostés sont de bons candidats. Ces dernières permettent de modéliser des relations très complexes et non linéaires entre un événement d'intérêt et les covariables. La littérature sur ce sujet est récente, plus particulièrement en ce qui concerne la mise en œuvre de ces méthodes sur de grandes bases de données de sinistres avec des événements potentiellement rares (Schneeweiss *et al*. 2017 ; Franklin *et al*. 2017).

Afin de développer un cadre méthodologique général pour les études pharmaco-épidémiologiques à partir des données du SNDS, ce projet s'appuie sur un ensemble d'associations de référence choisies *a priori*, cet ensemble devant être constitué de couples médicament-effets avec différents degrés de certitude quant à la plausibilité de leurs associations. Le principal événement d'intérêt est la prématurité, mais d'autres effets pourront également être considérés. Cet ensemble de référence permet l'évaluation et l'étalonnage des outils statistiques impliqués dans l'axe 2 du projet.

3.1 Axe 1

Dans le travail de thèse de Liliane Manitchoko, notre démarche a été d'étudier par simulations les performances des designs et modèles temporels considérés et de leurs techniques d'ajustement, dans un cadre qui est celui de données de cohorte. Nous sommes donc partis de données de cohorte simulées, pour ensuite sélectionner l'ensemble des "cas" ayant présenté l'évènement au cours du suivi et d'échantillonner des "témoins" parmi les individus à risque. Nous avons aussi étudié des méthodes d'analyse basées sur les cas seuls mais n'en présenterons pas ici les résultats. En effet, les résultats obtenus pour le schéma cas-croisé en particulier étaient décevants, voire contre-intuitifs : nous avons observé un biais qui augmentait avec le nombre de périodes témoins. Le temps a manqué pour explorer et comprendre d'où pouvaient provenir les biais observés. Nous avons donc préféré ne présenter ici que le travail abouti sur l'analyse cas-témoins nichée dans une cohorte.

3.1.1 Protocole de simulation

- **Génération des covariables fixes**

Nous avons utilisé des simulations pour étudier et comparer les propriétés statistiques des estimations obtenues au moyen des différentes méthodes d'analyse considérées. Pour chaque scénario considéré, nous avons généré 1000 échantillons aléatoires indépendants, chacun représentant une cohorte hypothétique de n = 5 000 sujets suivis pendant m = 2 ans (730 jours).

Nous avons considéré deux covariables fixes dans le temps (mesurées à l'inclusion) :

- une variable continue, l'âge ($A_i \sim$ Uniforme(40, 65)) en années compris entre 40 et 65 ans (à l'image de la cohorte E3N);
- une variable binaire, le sexe ($S_i \sim$ Bernoulli(0, 4)) avec 40% d'hommes comme considéré dans l'étude de simulation de Sylvestre et Abrahamowicz (2008).

- **Génération de l'exposition**

Le statut d'exposition Z_i d'un individu i, i = 1, ..., n a été généré comme une variable binaire (qui prend la valeur 1 lorsqu'un individu est exposé et la valeur 0 s'il ne l'est pas) à partir d'une distribution de Bernoulli avec une probabilité p d'exposition à l'inclusion pré-spécifiée et maintenue inchangée pendant toute la période de suivi. Cette proportion p pouvait dépendre des covariables fixes. En fonction de l'âge et du sexe, elle était définie comme :

$$\text{logit}(p_i) = \alpha_0 + \alpha_A \times A_i + \alpha_S \times S_i$$

où A_i est l'âge en années du sujet i, S_i est le sexe du sujet i avec $S_i = 1$ pour les hommes et $S_i = 0$ pour les femmes. Les paramètres de la fonction logistique $\alpha_A = \log(2)/10$ et $\alpha_S = \log(2)$ reflètent la dépendance de l'exposition à l'âge et au sexe tandis que le paramètre

$$\alpha_0 = \text{logit}(p) - (\log(2)/10 \times \text{moy}(A) + \log(2) \times \text{moy}(S))$$

peut être ajusté pour approcher une proportion voulue d'exposés dans la cohorte entière (moy(A) et moy(S) étant les valeurs moyennes des deux covariables fixes).

Exposition fixe dans le temps

Pour une exposition fixe dans le temps, les individus ayant un statut "exposé" ou "non exposé" sont considérés comme tels pendant toute la période d'observation.

Exposition variable dans le temps avec changement unique

Pour une exposition variable au cours du temps à changement unique, telle que l'initiation d'un traitement pendant le suivi, l'état d'exposition $Z_i(t)$ au temps t a été supposé dichotomique : le statut d'exposition des individus peut changer de "non exposé" à "exposé" au maximum une fois pendant l'intervalle de suivi. Une fois exposé, un sujet reste exposé pendant toute la durée restante du suivi.

Ainsi, pour chaque sujet exposé i identifié, nous avons généré son temps de début d'exposition $\tau(i)$ à partir d'une distribution uniforme sur l'intervalle [0;m] où m égale la durée maximale du suivi pour la cohorte entière. Ensuite, nous avons fixé $Z_i(t) = 0$ pour $0 < t < \tau(i)$ et $Z_i(t) = 1$ pour $t \geq \tau(i)$.

Exposition variable dans le temps avec changements multiples

Pour ce type d'exposition, nous avons distingué deux cas : celui d'une exposition ponctuelle répétée (vaccin à plusieurs doses par exemple) et celui d'une exposition répétée de durée variable (traitement antibiotique par exemple).

Dans le cas d'une exposition ponctuelle répétée, le statut d'exposition $Z_i(t)$ prend successivement les valeurs 0 ("non exposé") et 1 ("exposé") avec un premier changement au temps $t_1(i)$, un deuxième de 1 à 0 au temps $t_2(i)$, un troisième au temps $t_3(i)$, etc. Nous avons généré les temps de changement de statut d'exposition successifs de façon aléatoire :

- $t_1(i) \sim U(1, m)$: loi uniforme sur toute la période du suivi, $t_1(i)$ = début de la première période d'exposition, la durée d'exposition (d) étant fixée à 1 jour ;
- $t_2(i) \sim U[t_1(i) + d, m]$: loi uniforme entre la fin de la première période d'exposition et le dernier jour du suivi, $t_2(i)$ = début de la deuxième période d'exposition ;
- etc.

De cette façon, le nombre de changements de statut d'exposition était lui-même aléatoire (variable d'un individu à l'autre), comme dans l'étude de Hernán et al. qui évaluait l'association entre le vaccin recombinant contre l'hépatite B et le risque de sclérose en plaques : le nombre de fois que les individus avaient été vaccinés variait de 0 à 3 ou plus pendant la période considérée (Hernan *et al* 2004).

Dans le cas d'une exposition répétée de durée variable, la procédure de génération était similaire à celle du cas précédent mais avec une durée d'exposition aléatoire définie selon une distribution normale tronquée ($d \sim tN(m = 14, sd = 2, lo = 0, up = 28)$) de sorte que les valeurs soient comprises entre 0 et 28 jours et presque exclusivement dans l'intervalle [0; 18].

- **Génération des temps de censure**

Nous avons supposé une distribution uniforme pour générer les temps de censure supposés refléter les pertes de vue aléatoires au cours du suivi, avec une censure administrative supplémentaire de tous les sujets qui restaient à risque à l'issue des deux ans de suivi.

- **Génération des temps d'événement**

La simulation des temps d'évènement comme fonction de l'exposition fixe ou variable dans le temps a été effectuée à l'aide de l'algorithme de permutation spécialement développé et validé pour simuler les temps d'évènement conditionnellement aux effets et/ou covariables variant au cours du temps (Sylvestre et Abrahamowicz, 2008).

La fonction *permalgorithm* permettant d'appliquer cet algorithme est implémentée dans le package *PermAlgo* du logiciel R et nécessite le nombre d'individus n, la durée maximale de suivi pour ces individus m, une matrice contenant les valeurs possibles des covariables, le vecteur des coefficients de régression β correspondant à ces covariables, les temps d'évènement et de censure pour chaque individu. Pour nos simulations, nous avons supposé une loi exponentielle de paramètre λ constant (choisi selon la proportion d'exposés souhaitée) pour générer les temps d'évènement.

L'algorithme de permutation a été utilisé pour attribuer des temps d'évènement et de censure (classés par ordre croissant) aux vecteurs individuels des valeurs des covariables sur la base d'hypothèses préétablies concernant l'impact de ces covariables sur le risque de survenue d'évènement. Pour nos simulations, nous avons supposé le modèle de Cox défini par

$$h(t|Z, A, S) = h_0(t) \exp(\beta \times Z + \log(2)/10 \times A + \log(2) \times S)$$

ou

$$h(t|Z(t), A, S) = h_0(t) \exp(\beta \times Z(t) + \log(2)/10 \times A + \log(2) \times S)$$

pour une exposition fixe et variable dans le temps respectivement, où β désigne le paramètre associé à l'exposition d'intérêt.

Notre unité de temps pour l'analyse était d'un jour, de sorte que les temps d'évènement et de censure sont arrondis au jour le plus proche.

- **Scénarios de simulation**

Nous nous sommes inspirés d'exemples publiés (Austin *et al* 2012; Brickner 2015) pour le choix des scénarios considérés pour nos simulations. Pour chacun des deux types d'exposition fixe et variable dans le temps, nous avons supposé deux valeurs différentes pour l'effet théorique de l'exposition : β = log(2) = 0,6931 ou β = log(1,25) = 0,2231, soit un hazard ratio (HR) de 2 ou 1,25. Nous avons également fait varier les proportions de sujets exposés (0,10, 0,25 et 0,50) et de sujets ayant connu un évènement non censuré (0,05, 0,10 et 0,25) au cours des deux années de suivi. Cette dernière proportion a été approximativement contrôlée en ajustant le paramètre d'incidence λ de la distribution exponentielle utilisée pour générer les temps d'évènement. Dans chacun des 18 scénarios ainsi considérés pour chaque type d'exposition fixe ou variable dans le temps, les 1000 ensembles de données de cohorte simulées ont été analysés.

Ce même protocole de simulation a été utilisé pour étudier les propriétés statistiques des méthodes d'analyse basées sur les schémas de cohorte, cas-témoins niché et cas-croisé. Les résultats obtenus étant surprenants, nous avons choisi, pour mieux comprendre la provenance des biais observés, de nous recentrer sur les schémas de cohorte et cas-témoins niché avec des modèles simples sans covariables, comprenant uniquement une exposition fixe ou variable dans le temps à changement unique, soit ($h(t|Z) = h_0(t) \exp(\beta \times Z)$ ou $h(t|Z(t)) = h0(t) \exp(\beta \times Z(t))$) respectivement. Les résultats présentés ici sont donc restreints à ces schémas et types d'exposition.

3.1.2 Analyse statistique

Les données simulées ont été analysées à l'aide de deux approches : l'une correspondant à une étude de cohorte entière et l'autre à une étude cas-témoins nichée issue de la même cohorte.

- **Analyse des données de la cohorte entière**

Les données simulées ont été analysées à l'aide d'un modèle à risques proportionnels de Cox non ajusté avec une exposition Z fixe ou Z(t) variable dans le temps. La durée de suivi a été utilisée comme échelle de temps (Thiébaut & Bénichou 2004). Les estimations de l'effet (log-HR) de l'exposition, ainsi que l'intervalle de confiance (IC) à 95 %, ont été obtenus à l'aide de la fonction *coxph* du package *survival* de R, avec l'approximation d'Efron (1977) pour gérer les temps d'évènements ex-æquos. Nous nous sommes par ailleurs intéressés à l'approximation de Breslow (1974) en complément de celle d'Efron, car c'est l'option par défaut dans le logiciel SAS.

Nous avons quantifié formellement le nombre de temps d'évènements exæquos dans les jeux de données réelles et simulées à l'aide des formules suivantes. Soit M_j le nombre de temps d'évènements qui apparaissent j fois dans un ensemble de données. Le nombre total d'évènements dans cet ensemble de données M est défini par :

$$M = \text{Somme sur } j \geq 1 \text{ des } j \times M_j$$

et le nombre K de temps d'évènements distincts par :

$$K = \text{Somme sur } j \geq 1 \text{ des } M_j.$$

Le nombre de temps d'évènements ex-æquos est obtenu par soustraction du nombre K de temps d'évènements distincts au nombre total M d'évènements de la cohorte simulée :

$$M - K = \text{Somme sur } j \geq 1 \text{ des } (j - 1) \times M_j$$

- **Sélection et analyse des données cas-témoins nichées dans la cohorte**

A partir des données de cohorte simulées, des données pour une étude cas-témoins nichée ont été échantillonnées. Pour les données simulées, à chacun des 1000 jeux de données correspond une étude cas-témoins.

L'échantillonnage des données cas-témoins nichées a été effectué en identifiant d'abord tous les "cas" (sujets qui ont eu l'évènement pendant le suivi) et leur temps d'évènement correspondant dans la cohorte, puis en sélectionnant aléatoirement par densité d'incidence (dans l'ensemble à risque de chaque cas considéré) 1 ou 5 témoins à apparier à chaque cas. Le choix de ces deux scénarios du nombre de témoins par cas s'est fait en s'inspirant de celui proposé par Austin (2012). Pour l'appariement 5:1, les témoins ont été sélectionnés de manière imbriquée, de sorte que le témoin sélectionné pour l'appariement 1:1 faisait partie des cinq témoins sélectionnés. Cette procédure d'échantillonnage a été effectuée en utilisant la fonction *ccwc* du package *Epi* de R.

Les échantillons cas-témoins nichés avec 1 ou 5 témoins par cas ont été analysés à l'aide du modèle de régression logistique conditionnelle pour estimer l'association entre l'exposition et le risque de survenue de l'évènement d'intérêt, en conditionnant sur les ensembles appariés. Les estimations de l'effet (log-Odds Ratio) de l'exposition pour la régression logistique conditionnelle, équivalentes au log-HR, ont été obtenues à l'aide de la fonction *clogit* du package *survival* de R, avec l'approximation d'Efron pour gérer les ex-æquos au sein des strates.

- **Exploration des biais dans l'analyse des données cas-témoins nichées**

Biais dus aux données éparses

Pour explorer la possibilité de biais dus aux données éparses, nous avons réanalysé les échantillons cas-témoins nichés en utilisant la régression logistique conditionnelle avec la vraisemblance pénalisée de Firth (1993). Nous avons utilisé le code R d'abord développé par Sun *et al.* (2011) pour étendre l'approche de réduction de biais de Firth à la régression logistique conditionnelle, puis étendu par Cheung *et al*. (2019) pour un nombre illimité de variables d'exposition.

Biais dus aux temps d'événements ex-aequos

Nous avons ensuite exploré la possibilité de biais dus à la gestion des temps d'évènements ex-aequos selon deux stratégies :

1- Modification de la méthode de gestion des ex-æquos

Pour explorer l'impact du choix d'une méthode de gestion des temps d'évènements ex-æquos sur les estimations obtenues par les analyses de régression logistique conditionnelle des études cas-témoins nichées, nous avons refait les analyses en considérant deux méthodes alternatives à l'approximation d'Efron : la méthode exacte (Peto 1972) et l'approximation de Breslow.

2- Modification de la sélection des cas et témoins appariés

La fonction *ccwc* que nous avons utilisée pour la sélection des témoins appariés à chaque cas regroupe tous les cas qui ont subi l'évènement d'intérêt au même moment ainsi que leurs témoins respectifs dans la même strate, ce qui peut entraîner la présence d'évènements ex-æquos dans les strates lorsque deux cas ou plus ont subi l'évènement au même moment.

En nous inspirant de la procédure d'échantillonnage des données cas-témoins nichées dans une cohorte de la macro SAS %nCCsampling (Desai *et al* 2016), nous avons trouvé un moyen de nous affranchir des problèmes liés à la présence d'ex-æquos dans les strates. En effet, nous avons modifié le code de la fonction *ccwc* en utilisant un identifiant unique pour chaque cas de façon à ce que les cas qui ont vécu l'évènement au même moment soient alloués à des strates distinctes. Ainsi nous avons veillé à ce que chaque strate ne comprenne qu'un seul cas, ce qui permet d'éviter tout ex-æquos. Le code modifié de la fonction *ccwc* est présenté en matériel supplémentaire de l'article publié (Desai *et al* 2016). Les échantillons cas-témoins nichés sélectionnés à l'aide de cette fonction R modifiée ont été

analysés par régression logistique conditionnelle avec l'approximation d'Efron, les trois méthodes de gestion des ex-æquos (l'approximation de Breslow et la méthode exacte) aboutissant au même résultat étant donné l'absence d'ex-æquos dans les strates. Cette approche a été appelée "approche *ccwc* modifiée".

- **Critères de comparaison pour les données simulées**

Pour chaque scénario considéré, nous avons comparé les performances des méthodes statistiques utilisées pour analyser les données cas-témoins nichées et celles de la cohorte entière en utilisant les critères suivants :

- le biais relatif par rapport à la valeur théorique de β utilisée dans la procédure de génération des données
- la moyenne de tous les écarts-types estimés sur l'ensemble des simulations k = 1, ..., 1000 (SEE pour *average standard error estimator*)
- l'erreur quadratique moyenne (RMSE pour *Root Mean Squared Error*)
- la probabilité de couverture (ou taux de recouvrement) : proportion d'échantillons simulés pour lesquels l'intervalle de confiance à 95% de l'effet estimé contient la valeur théorique β. Pour 1000 échantillons simulés, on s'attend à un taux de recouvrement compris entre 0,9365 et 0,9635.

Dans cet axe, nous nous intéressons aux méthodes et modèles statistiques de prise en compte de la confusion mesurée et non mesurée et nous les mettons en œuvre pour évaluer des associations médicamenteuses pendant la grossesse sur les données du SNDS. Cet axe est constitué de trois volets : la constitution d'un ensemble de référence d'associations médicamenteuses pour la grossesse, la mise en œuvre de scores synthétiques en grande dimension, appliquée à la prématurité, la mise en œuvre « grandeur réelle » d'une analyse par variable instrumentale d'une association de référence. Les deux premiers volets sont étroitement liés, et abordés avec l'échantillon permanent des bénéficiaires du SNDS (EGB) ; le troisième s'appuie sur une association de référence obtenue au cours de cette recherche (1ère volet) pour l'évaluer sur les données du SNDS complètes.

3.2.1 Établissement d'un ensemble d'associations labellisées d'intérêt pour la grossesse

Un première étape du projet prévue sur les 6 premiers mois était de construire un large ensemble de référence constitué d'associations médicament – événement indésirable avec différents degrés de certitude quant à la plausibilité de leur association, en tout premier lieu pour la prématurité : des associations bien documentées (contrôles positifs), des associations pour lesquelles il n'y a aucune preuve disponible pouvant suggérer un effet causal du médicament (contrôles négatifs), et des associations actuellement discutées. Il était prévu que cette construction se fasse sur la base d'un panel

d'experts cliniciens et pharmacologues, ainsi que de la littérature. Nous avions obtenu l'accord de ces experts pour « balayer », les pharmacologues, puis les cliniciens, un grand nombre d'associations potentielles que nous leur avions fournies, constituées des médicaments suffisamment prescrits chez les femmes enceintes dans les données de l'EGB (échantillon au 1/97ème du SNDS), sans autre pré-sélection, et pour vingt pathologies de grossesse, dont la prématurité. Malgré leur bonne volonté et intérêt, les experts pharmacologues n'ont pas pu dans leur majorité effectuer ce travail de catégorisation, qui s'est avéré d'une ampleur incompatible avec les crises sanitaires du moment qu'ils avaient à gérer. Nous avons alors modifié la procédure de construction.

Les médicaments inducteurs de la prématurité sont peu connus et il n'existe aucune base de données permettant de les recenser. Afin d'alléger l'étape de l'expertise, nous avons tout d'abord effectué une pré-sélection des médicaments, à l'aide de méthodes de détection de signaux (présentées dans la section suivante) que nous avons évaluées et développées, particulièrement adaptées pour les bases de données médico-administratives. Nous utilisons 21 méthodes différentes de détection de signaux pour les grossesses d'une durée supérieure à 22 semaines d'aménorrhée (terme minimal pour définir un accouchement prématuré) identifiées dans l'EGB entre 2009 et 2016. Nous avons retenu les médicaments qui ont été prescrits au cours d'au moins 5 grossesses. Nous avons choisi d'étudier l'effet des expositions médicamenteuses durant un dernier trimestre d'une durée fixe calculée à partir de la date d'accouchement (T_{BD} pour « trimester before delivery »). Les signaux sont ensuite classés à partir de l'évaluation clinique de la littérature scientifique publiée et de l'avis d'une experte pharmacovigilante entre les 3 modalités : « Aberrant », « Potentiel » et « Incertain ». Par rapport à la catégorisation initiale, les associations classées en « Potentiel » incluent à la fois les contrôles positifs et ceux qui sont en discussion, tandis que les « Aberrantes » constituent les contrôles négatifs.

Bien qu'efficace, cette stratégie, qui repose sur l'expertise malgré tout, reste lourde si on veut l'élargir à d'autres pathologies au-delà de la prématurité. Nous avons donc développé un requêtage automatisé à partir de MEDLINE pour 20 pathologies dont 13 spécifiques de la grossesse, et 7 plus générales : i) Fausse Couche Spontanée (ii) Grossesse Extra-Utérine (iii) Prématurité (iv) Mort Fœtale In Utero (v) Retard de Croissance Intra Utérin (vi) Macrosomie (vii) Oligohydramnios (viii) Polyhydramnios (ix) Rupture Prématurée des Membranes (x) Hypertension Artérielle Gravidique et ses formes compliquées (xi) Diabète Gestationnel (xii) Cholestase gravidique (xiii) Vomissements (xiv) Hémorragies (xv) Thromboses (xvi) Infections (xvii) Colique néphrétique (xviii) Pancréatite (xix) Insuffisance rénale aiguë (xx) Insuffisance hépatique aiguë. Nous avons ensuite proposé une méthode pour quantifier l'état des connaissances scientifiques d'une association donnée qui puisse permettre de l'annoter.

3.2.2 Mise en œuvre et comparaison des scores synthétiques en grande dimension pour la prise en compte de la confusion : application à la prématurité

L'utilisation du score de propension (PS) est devenue courante en pharmacoépidémiologie afin de prendre en compte la confusion mesurée. Plus récent, le score de propension en grande dimension hdPS exploite la grande richesse des informations dans les bases de données médico-administratives (Schneeweiss *et al* 2009). Le hdPS consiste à présélectionner de manière automatisée les variables entrant dans la construction du score suivant leurs association avec l'événement et le traitement d'intérêt. Certains auteurs ont émis l'hypothèse qu'en exploitant un grand nombre de variables on pourrait indirectement décrire l'état de santé des patients et ainsi capturer également de la confusion non mesurée. Quelle que soit la méthode pour l'estimer (grande dimension ou pas), l'analyse consiste ensuite à apparier, ajuster, pondérer ou stratifier sur cette prédiction d'être traité par le médicament d'intérêt. Nous nous sommes aussi intéressés à un autre score, le score pronostique (DRS, Hansen 2008) qui est une alternative performante lorsque l'exposition est rare. Le modèle de prédiction est établi chez les non exposés. L'intégration du score pronostique se fait par appariement, ajustement ou stratification.

Les données utilisées sont celles décrites dans la section précédente. Nous avons choisi d'étudier l'effet des expositions médicamenteuses sur la prématurité durant T_{BD} et d'utiliser les expositions à T_0 (un trimestre avant la conception) et à T_1 (premier trimestre de grossesse) comme variables d'ajustement dans le modèle final. Les autres variables utilisées étaient des variables socio-démographiques (âge maternel, statut CMU-C, indice de défavorisation sociale FDep), ainsi que de nombreuses pathologies.

Nous avons voulu tester l'importance des choix méthodologiques pour la construction des scores de propension et pronostiques et leur intégration dans l'analyse. Le premier concerne le choix des variables à inclure dans le modèle avec d'un côté un algorithme sur données complètes et de l'autre un algorithme basé sur une présélection de variables. Pour les deux scores, cette présélection a été faite grâce à la formule de Bross, issue de la procédure hdPS, implémentée dans le package R « hdps ». A partir des expositions médicamenteuses par trimestre, des pathologies associées et des variables socio-démographiques, nous avons choisi de présélectionner 200 variables à inclure dans la construction des modèles des scores, en plus des 3 variables sociodémographiques. Nous avions ainsi deux modalités pour la construction des scores, d'un côté une base de variables non filtrées (1222 variables) et de l'autre une base de 203 variables. Le deuxième choix concerne les méthodes d'apprentissage automatique pour gérer un grand nombre de variables. Nous avons ici évalué deux méthodes adaptées à la sélection de variables en grande dimension : la régression logistique lasso (package « glmnet ») et les arbres de régression boostés (boosting, package « xgboost »). Afin de rendre les deux analyses les plus comparables, nous avons utilisé la même partition de notre population, pour la validation croisée entre le lasso et le boosting. Le troisième choix concerne la façon d'intégrer le score dans l'analyse. Nous avons évalué l'appariement, l'ajustement et la pondération pour PS, et l'appariement et l'ajustement pour DRS. Nous avons également ajusté un modèle brut sans score, ce qui fait au total 21 méthodes, avec 8 combinaisons pour DRS et 12 combinaisons pour PS. L'ensemble des calculs ont été réalisés sur un serveur de calculs à l'aide du logiciel R.

3.3.3 Prise en compte de la confusion non mesurée par l'utilisation de variables instrumentales

Nous avons pu aborder le dernier volet de ce 2ème axe portant sur les méthodes basées sur les variables instrumentales pour corriger le biais de confusion des estimations de risque, lorsque certains facteurs de confusion ne sont pas mesurés. Nous mettons en œuvre en « grandeur réelle », sur les données médico-administratives du SNDS, les méthodologies adéquates. Cette technique est encore peu utilisée en pharmacoépidémiologie, notamment parce qu'elle nécessite une mise en œuvre rigoureuse d'un point de vue statistique. Une autre difficulté pour ce type d'approche est de définir les bons comparateurs en termes de classes de médicaments. Notre investigation porte sur les inhibiteurs sélectifs de la recapture de la sérotonine (ISRS) qui ont été parfois décrits dans la littérature comme pouvant être associés à des issues pathologiques de la grossesse, notamment la prématurité (Eke *et al* 2016).

- **Population d'intérêt**

La population étudiée est composée de toutes les femmes enceintes françaises enregistrées dans la base de données SNIIRAM et ayant accouché entre 2012 et 2014. Plus précisément, environ 2,1 millions de grossesses ont été identifiées à partir des codes de sortie d'hôpital. Seules les grossesses supérieures ou égales à 22 semaines ont été incluses dans cette extraction. Par conséquent, les grossesses plus courtes ne sont pas disponibles dans notre base de données.

- **Choix de l'association d'intérêt et covariables**

Nous étudions le risque de prématurité pour une grossesse exposée à l'escitalopram (code ATC N06AB10) de la classe des ISRS (code ATC N06AB). Cette association est l'un des signaux que nous avons obtenus. Nous reprenons donc les mêmes définitions :

On définit la prématurité comme une naissance vivante dans [22 SA – 37 SA[et la fenêtre d'exposition d'intérêt est T_{BD}. Le comparateur est les autres antidépresseurs de la classe ISRS. Les covariables âge de la mère, variables sociodémographiques, facteurs de risque de la prématurité sont prises en compte, ainsi que l'exposition à T_1 (suite aux analyses faites pour la constitution de l'ensemble de référence). Nous explorons également la spécialité du prescripteur.

- **Rappel sur la méthode des variables instrumentales**

Dans les études d'observation, les facteurs de confusion non observés peuvent fausser l'estimation de l'effet. Au cours de la dernière décennie, cette question a fait l'objet d'une attention croissante dans le domaine des études épidémiologiques visant à évaluer les effets indésirables des médicaments, et quelques travaux se sont concentrés sur les approches basées sur les variables instrumentales (VI).

L'estimation des variables instrumentales est une approche bien connue pour évaluer l'endogénéité dans la modélisation statistique (Rassen *et al* 2009). L'endogénéité survient souvent lorsqu'un modèle causal est mal spécifié, ce qui introduit un biais structurel dans l'estimation de ses paramètres. Cela peut résulter d'une erreur de mesure des variables (Caroll *et al* 1995), d'une variable non observée (Nestler 2015) ou d'une causalité inverse entre la réponse et certains régresseurs. La méthode générale d'estimation IV tente d'éliminer ce biais en utilisant des équations structurelles qui intègrent des variables instrumentales dans le modèle. Plusieurs approches théoriques ont été développées pour construire des estimateurs de paramètres et étudier les propriétés de ces estimateurs dans les modèles causaux avec endogénéité (voir Greene 2012) [5]). Un exemple bien connu est le cas des modèles linéaires dans lesquels l'estimation VI conduit à des estimateurs présentant de bonnes propriétés de convergence. Le biais structurel peut être complètement éliminé dans ce cas.

On considère le modèle logistique liant une réponse binaire *Y* (événement / pas événement) à une exposition binaire *T* (T_1/T_2) :

$$\mathbb{E}(Y|T,X,X_u) = expit(\gamma_0 + T\gamma + X\gamma_X + U\gamma_U)$$

où *X* est une covariable et *U* une variable de confusion non observée. L'objectif est d'estimer le paramètre γ. Si on applique aux données observées le modèle logistique conventionnel sans *U*, l'estimation de l'effet de *T* est biaisée. Le biais peut être réduit en utilisant un instrument valide, c'est-à-dire une variable indépendante de la variable de confusion non mesurée, corrélée avec l'exposition (force de l'instrument), sans association directe avec l'événement d'intérêt. Ces hypothèses ne sont pas toutes testables en pratique. Plusieurs méthodes d'analyse par variable instrumentale ont été proposées. Dans une étude de simulation, nous les avions comparées dans de nombreux scénarios avec différents niveaux de confusion et différentes forces d'instruments (Koladjo *et al* 2018), et observé que la méthode 2SRI (Two-Stage Residual Inclusion) était la moins biaisée. Cette procédure d'estimation en deux étapes consiste à ajuster une première régression logistique liant *T* et l'instrument *Z*

$$\mathbb{E}(T|Z,X) = expit(\alpha_0 + Z\alpha + X\alpha_X)$$

puis à effectuer une deuxième régression logistique en incluant les résidus *R* de la première régression dans le modèle

$$\mathbb{E}(Y|T,R,X) = expit(\beta_0 + T\beta + X\beta_X + R\beta_R)$$

pour estimer β. A noter que la variance de cette estimation n'est pas celle issue de de la 2ème régression et qu'il faut la corriger pour prendre en compte la procédure en deux étapes (la formule est dans l'article de Koladjo *et al* 2018), ou bien appliquer une procédure bootstrap. Dans cette même étude, il avait été montré que le bootstrap ou la correction analytique de la variance donnaient des résultats semblables. Plus généralement, Ertefaie *et al* (2017) propose un tutorial pour l'utilisation des VI en pharmacoépidémiologie. Dans ce travail, nous appliquons donc la méthode ESRI, avec les deux manières de calculer l'intervalle de confiance (1000 bootstraps).

- **Choix et construction de la variable instrumentale**

Les variables instrumentales les plus souvent utilisées en pharmaco-épidémiologie sont basées sur les préférences des médecins en matière de prescription. Il s'agit de comparer les deux traitements T_1 et T_2 lorsque l'attribution du traitement au patient n'est pas randomisée. Brookhart et *al* (2006) proposent d'utiliser le traitement prescrit par un médecin à son patient précédent comme proxy de cette VI pour son nouveau patient. Dans Koladjo *et al* (2018), le comportement de Z = traitement précédent a été étudié dans le contexte de la régression logistique. Il a également été comparé à celui de la VI basée sur la préférence du médecin sous la forme continue c'est-à-dire la proportion de tous les patients précédents du même médecin qui se sont vus prescrire le traitement en question. Cela correspond à l'estimateur empirique de la probabilité pour un médecin de préférer le traitement en question. Dans ce travail, c'est ce que nous avons utilisé.

Les estimations sont réalisées avec les fonctions de régression existantes (glm) implémentées dans le logiciel statistique R.

4. Résultats

4.1 Axe 1

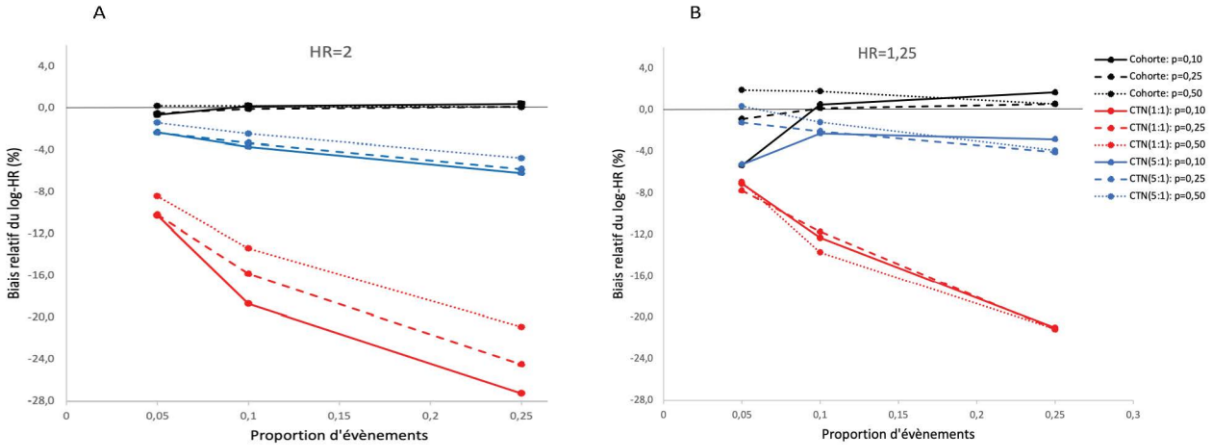
La thèse de Liliane Manitchoko s'est concentrée sur l'analyse cas témoins nichée. En effet, si beaucoup de travaux existent sur la comparaison de ce schéma d'étude avec l'analyse de la cohorte entière, ceux-ci concernent pour une très grande majorité une exposition fixe dans le temps. Un seul article s'était intéressé spécifiquement à une exposition variable dans le temps et montrait des résultats de simulation contre-intuitifs, à savoir que les estimations issues d'une analyse cas-témoins nichée pouvaient être biaisées, le biais augmentant avec la proportion d'évènements (Austin *et al* 2012). Dans un premier temps, nous avons analysé les données simulées à l'aide des procédures disponibles dans le logiciel R (section 4.1.1). Ayant constaté des biais similaires à ceux rapportés précédemment, nous avons cherché à corriger ces biais à l'aide d'une méthode de réduction de biais (section 4.1.2). Nous avons ensuite recherché la cause de ces biais en lien avec la prise en compte des temps d'évènement ex-æquos (section 4.1.3).

4.1.1 Analyses de la cohorte entière et cas-témoins nichées (approximation d'Efron)

- **Exposition fixe dans le temps**

La figure 1A représente les biais relatifs des estimations du log-HR obtenues à partir des analyses de la cohorte entière et des données CTN pour une exposition fixe dans le temps et un HR théorique de 2.

Figure 1 – *Biais relatif du log-HR estimé dans les schémas de cohorte et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition fixe dans le temps et hazard ratio théoriques de 2 (A) et 1,25 (B)*



Le tableau 1 contenant l'ensemble des résultats de simulation complète ces résultats avec la moyenne des écarts-types estimés, l'erreur quadratique moyenne et la probabilité de couverture. Dans les analyses de la cohorte entière comme CTN, l'approximation d'Efron est l'approche utilisée par défaut pour la gestion des ex-æquos.

Tableau 1 – *Résultats des analyses de la cohorte entière et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition fixe dans le temps et hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-0,636	0,167	0,166	95,3	0,200	0,116	0,113	95,5	0,406	0,076	0,076	95,4
CTN (1:1)	-10,256	0,244	0,251	93,8	-18,677	0,155	0,194	87,6	-27,263	0,091	0,205	45,7
CTN (5:1)	-2,311	0,188	0,184	95,0	-3,673	0,128	0,128	95,4	-6,190	0,081	0,092	91,6
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-0,498	0,130	0,129	95,4	-0,076	0,090	0,089	95,8	0,137	0,058	0,057	96,0
CTN (1:1)	-10,193	0,182	0,187	94,0	-15,829	0,117	0,154	86,1	-24,480	0,068	0,181	28,4
CTN (5:1)	-2,361	0,143	0,145	95,4	-3,328	0,098	0,098	95,0	-5,812	0,061	0,073	91,0
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	0,207	0,135	0,137	94,8	0,235	0,093	0,091	94,8	0,089	0,059	0,059	95,5
CTN (1:1)	-8,408	0,177	0,187	93,9	-13,388	0,116	0,143	88,8	-20,890	0,068	0,158	42,1
CTN (5:1)	-1,397	0,144	0,147	94,5	-2,446	0,098	0,097	95,4	-4,752	0,061	0,070	91,5

SEE *average standard error estimator*; RMSE *root mean squared error*; PC probabilité de couverture; CTN cas-témoins niché

Dans tous les scénarios, les estimations du log-HR pour l'analyse de la cohorte entière étaient non biaisées (biais relatif <1%par rapport au log-HR théorique). Les estimations du log-HR pour les analyses des données CTN présentaient un biais négatif qui diminuait avec l'augmentation du nombre de témoins par cas (biais relatif maximal de 27% pour un témoin par cas à 6% pour cinq témoins par cas, figure 1A et tableau 1). Ce biais avait tendance à augmenter avec la proportion d'évènements (pour 10% de sujets

exposés et un témoin par cas, le biais relatif passait de 10% à 27% pour 5% et 25% d'évènements respectivement).

Les estimations issues de l'analyse de la cohorte entière présentaient une variance plus faible que celles des analyses CTN quel que soit le scénario (voir les colonnes SEE présentant les moyennes des écart-types estimés dans le tableau 1). La précision des estimations s'améliorait de façon générale pour les deux schémas d'étude à mesure que la proportion d'évènements, la prévalence de l'exposition et (pour les analyses CTN) le nombre de témoins par cas augmentaient.

Dans tous les scénarios, des probabilités de couverture proches de la valeur nominale (0,95) ont été observées pour le schéma de cohorte tandis que les probabilités de couverture pour le schéma d'étude CTN étaient inférieures à la valeur nominale, en particulier pour des proportions d'évènements plus élevées (scénarios pour lesquels des biais considérables ont été observés).

Des résultats similaires à ceux présentés ci-dessus avec un HR=2 ont été obtenus pour des simulations avec un HR=1,25 (figure 3.1 B, tableau 3.2). Dans tous les scénarios, les estimations du log-HR pour l'analyse de la cohorte entière étaient non biaisées sauf pour le scénario de faibles proportions de sujets exposés et d'évènements (respectivement 10% et 5 %) pour lequel un léger biais négatif de 5% a été observé (figure 3.1 B). Un biais négatif (mais d'ampleur réduite par rapport à celui observé pour un HR=2) augmentant avec la proportion d'évènements (pour 10% de sujets exposés et un témoin par cas, le biais relatif passait de 7% à 21% pour 5% et 25% d'évènements respectivement) a également été observé dans les analyses CTN (figure 1B, tableau 2). Cependant, le biais étant plus faible, les probabilités de couverture étaient plus proches de la valeur nominale avec un HR=1,25 qu'avec un HR=2.

Tableau 2 – *Résultats des analyses de la cohorte entière et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition fixe dans le temps et hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-5,331	0,196	0,204	94,6	0,529	0,135	0,136	94,3	1,697	0,087	0,087	94,7
CTN (1:1)	-7,078	0,269	0,267	95,7	-12,308	0,173	0,169	95,8	-21,038	0,102	0,111	93,2
CTN (5:1)	-5,243	0,215	0,225	94,0	-2,251	0,145	0,144	95,6	-2,824	0,091	0,091	95,2
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-0,865	0,140	0,139	94,8	0,166	0,097	0,099	94,3	0,562	0,062	0,062	96,4
CTN (1:1)	-7,751	0,189	0,188	94,9	-11,707	0,123	0,126	94,3	-21,183	0,072	0,085	91,1
CTN (5:1)	-1,206	0,152	0,147	96,1	-2,046	0,104	0,105	95,0	-4,071	0,065	0,065	95,4
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	1,921	0,128	0,128	95,4	1,795	0,088	0,087	95,0	0,579	0,057	0,057	95,0
CTN (1:1)	-6,928	0,169	0,165	94,7	-13,706	0,111	0,111	94,5	-21,156	0,065	0,080	88,3
CTN (5:1)	0,356	0,138	0,139	94,5	-1,175	0,094	0,092	95,2	-3,902	0,059	0,060	94,7

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché

Figure 3 – *Biais relatif du log-HR des analyses CTN (pour 1 et 5 témoins par cas) avec les approximations de Breslow (A et B respectivement) et d'Efron (C et D respectivement) et la méthode exacte (E et F respectivement) : exposition fixe dans le temps et hazard ratio théorique de 2*

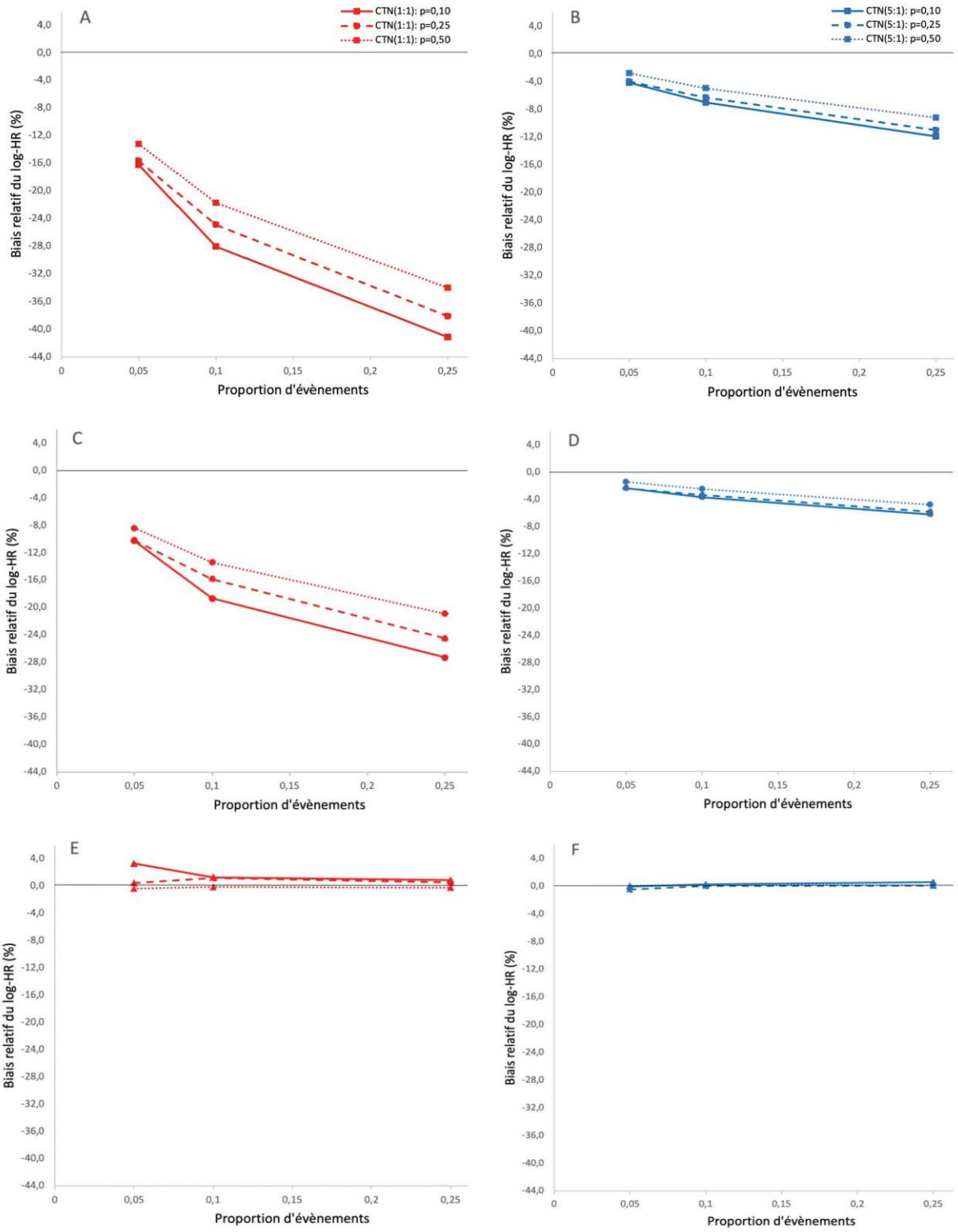


Tableau 3 – *Résultats des analyses de la cohorte entière et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition variable dans le temps et hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-2,328	0,226	0,226	96,1	-0,611	0,157	0,146	96,9	0,506	0,104	0,100	96,7
CTN (1:1)	-12,283	0,335	0,328	95,2	-19,010	0,213	0,232	92,4	-27,395	0,125	0,218	70,6
CTN (5:1)	-4,591	0,256	0,264	94,1	-4,080	0,174	0,165	96,5	-6,716	0,111	0,116	95,3
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-1,616	0,163	0,170	94,2	0,183	0,113	0,113	95,1	-0,820	0,074	0,074	94,5
CTN (1:1)	-12,551	0,234	0,251	93,0	-17,568	0,150	0,186	88,3	-26,514	0,088	0,200	45,9
CTN (5:1)	-3,458	0,182	0,191	94,3	-3,471	0,124	0,125	94,9	-7,119	0,079	0,091	90,8
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-1,766	0,143	0,147	94,0	-0,938	0,099	0,099	94,3	-0,044	0,064	0,066	94,6
CTN (1:1)	-11,772	0,198	0,211	92,6	-17,178	0,128	0,171	85,7	-24,454	0,075	0,184	37,3
CTN (5:1)	-3,611	0,157	0,160	93,8	-4,305	0,107	0,111	94,1	-5,983	0,067	0,081	89,2

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché

- **Exposition variable dans le temps**

Des résultats similaires à ceux d'une exposition fixe dans le temps ont été obtenus pour une exposition variable dans le temps avec un HR=2 (tableau 3) ou un HR=1,25 (tableau 4).

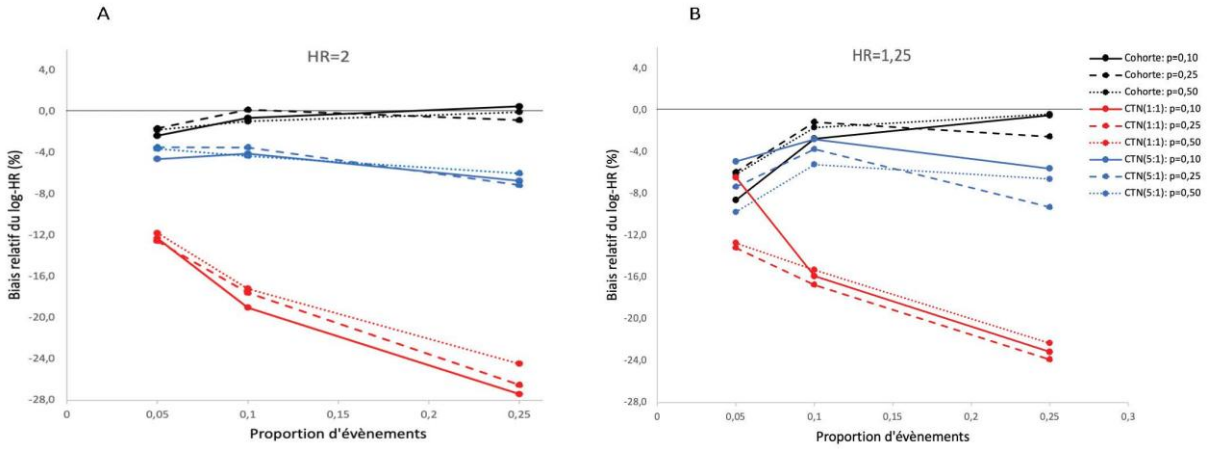
Tableau 4 – *Résultats des analyses de la cohorte entière et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition variable dans le temps et hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-8,613	0,273	0,287	95,2	-2,719	0,188	0,177	97,1	-0,470	0,123	0,123	95,3
CTN (1:1)	-6,424	0,376	0,382	95,6	-15,887	0,242	0,234	96,4	-23,160	0,144	0,145	95,4
CTN (5:1)	-4,907	0,299	0,304	95,3	-2,774	0,203	0,195	96,5	-5,576	0,129	0,131	94,7
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-5,921	0,185	0,184	95,7	-1,122	0,128	0,129	95,5	-2,531	0,084	0,085	94,3
CTN (1:1)	-13,185	0,252	0,250	95,1	-16,726	0,164	0,164	94,9	-23,891	0,098	0,110	92,0
CTN (5:1)	-7,342	0,202	0,206	95,4	-3,718	0,138	0,140	95,7	-9,278	0,088	0,091	94,6
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte	-6,144	0,151	0,158	94,7	-1,637	0,104	0,105	94,8	-0,359	0,068	0,071	93,9
CTN (1:1)	-12,709	0,203	0,208	94,3	-15,293	0,132	0,140	93,7	-22,324	0,079	0,094	91,5
CTN (5:1)	-9,748	0,163	0,171	94,4	-5,197	0,112	0,112	94,5	-6,585	0,071	0,075	94,8

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché

Dans tous les scénarios, l'analyse de la cohorte entière a donné des estimations non biaisées (biais relatif <2,5 %). Cependant, les estimations du log-HR issues des analyses CTN présentaient toujours un biais négatif qui diminuait avec l'augmentation du nombre de témoins par cas (biais relatif maximal de 27% pour un témoin par cas à 7% pour cinq témoins par cas) mais augmentait avec la proportion élevée d'évènements (figure 2A).

Figure 2 – *Biais relatif du log-HR estimé dans les schémas de cohorte et CTN (pour 1 et 5 témoins par cas) en fonction de la proportion d'évènements et de la prévalence de l'exposition p, les temps ex-æquos étant gérés selon l'approximation d'Efron : exposition variable dans le temps et hazard ratio théoriques de 2 (A) et 1,25 (B)*



Dans tous les scénarios, les variances des estimations issues de l'analyse de la cohorte entière étaient plus faibles que celles des estimations des analyses CTN (SEE dans le tableau 3). Une amélioration de la précision des estimations a été observée pour les analyses de données des deux schémas d'étude à mesure que la proportion d'évènements, la prévalence de l'exposition et (pour les analyses CTN) le nombre de témoins par cas augmentaient.

Des probabilités de couverture proches de la valeur nominale (0,95) ont été observées pour le schéma de cohorte dans tous les scénarios alors que certaines probabilités de couverture du schéma d'étude CTN étaient inférieures à la valeur nominale, en particulier pour des proportions d'évènements plus élevées (10% d'évènements pour un témoin par cas et 25% d'évènements pour 1 et 5 témoins par cas, tableau 3).

Les résultats des simulations avec un HR=1,25 montrent que les estimations du log-HR pour l'analyse de la cohorte entière étaient non biaisées sauf pour les scénarios avec 5% d'évènements. En effet, pour ces scénarios, on a observé un biais négatif compris entre 6% et 8% (figure 2B, tableau 4). De même que pour un HR=2, nous avons observé, pour un HR=1,25, un biais négatif (mais d'ampleur réduite par rapport à celui observé pour un HR=2) qui augmentait avec la proportion d'évènements (pour 25% de sujets exposés et un témoin par cas, le biais relatif passait de 13% à 24% pour 5% et 25% d'évènements respectivement) dans les analyses CTN (figure 2B„ tableau 4).

4.1.2 Correction des biais dus aux données éparses dans les analyses cas-témoins nichées par une méthode de réduction des biais

- **Exposition fixe dans le temps**

L'ensemble des résultats de simulation sont présentés dans le tableau 5. Dans tous les scénarios, nous avons constaté une diminution considérable des biais après recours à la méthode de Firth pour la réduction des biais. En effet, le biais relatif maximal était de seulement 5%pour les analyses CTN après

réduction des biais tandis qu'il était de 27% pour les mêmes analyses utilisant l'approximation d'Efron (tableau 5).

Tableau 5 – *Résultats des analyses CTN (pour 1 et 5 témoins par cas) avec réduction des biais selon la méthode de Firth : exposition fixe dans le temps et hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-10,256	0,244	0,251	93,8	-18,677	0,155	0,194	87,6	-27,263	0,091	0,205	45,7
CTN (1:1) RB	2,831	0,276	0,302	93,7	2,058	0,192	0,202	94,2	2,626	0,127	0,134	93,6
CTN (5:1) Ef	-2,311	0,188	0,184	95,0	-3,673	0,128	0,128	95,4	-6,190	0,081	0,092	91,6
CTN (5:1) RB	0,882	0,191	0,188	95,3	0,876	0,134	0,132	95,1	1,221	0,088	0,091	94,8
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-10,193	0,182	0,187	94,0	-15,829	0,117	0,154	86,1	-24,480	0,068	0,181	28,4
CTN (1:1) RB	1,413	0,203	0,210	94,4	3,738	0,142	0,151	94,0	4,595	0,093	0,105	90,9
CTN (5:1) Ef	-2,361	0,143	0,145	95,4	-3,328	0,098	0,098	95,0	-5,812	0,061	0,073	91,0
CTN (5:1) RB	-0,152	0,146	0,148	94,8	0,451	0,101	0,100	95,7	0,872	0,066	0,067	94,0
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-8,408	0,177	0,187	93,9	-13,388	0,116	0,143	88,8	-20,890	0,068	0,158	42,1
CTN (1:1) RB	0,849	0,191	0,205	94,8	2,784	0,133	0,141	93,9	4,432	0,086	0,099	91,2
CTN (5:1) Ef	-1,397	0,144	0,147	94,5	-2,446	0,098	0,097	95,4	-4,752	0,061	0,070	91,5
CTN (5:1) RB	-0,157	0,146	0,149	94,0	0,392	0,101	0,098	95,8	0,772	0,064	0,066	94,2

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture; CTN cas-témoins niché; Ef Efron; RB réduction de biais; les résultats des lignes grisées sont les mêmes que dans le tableau 1 et sont inclus ici pour faciliter la comparaison.

La précision des estimations après réduction de biais était systématiquement inférieure à celle des estimations des mêmes analyses CTN avant réduction des biais. Ainsi, l'erreur quadratique moyenne (RMSE) n'était avantageuse pour la méthode de réduction de biais que pour les scénarios avec des proportions d'évènements plus élevées (25 %) pour lesquels des biais considérables avaient été observés dans un premier temps (voir les colonnes RMSE dans le tableau 5).

Les très faibles probabilités de couverture observées précédemment pour le schéma d'étude CTN avec l'approximation d'Efron se sont nettement améliorées après réduction des biais (tableau 5).

Des résultats similaires ont été observés pour un HR théorique égal à 1,25. Une diminution considérable des biais a été observée après réduction des biais selon la méthode de Firth (biais maximal de 4%alors qu'il était de 21%avec l'approximation d'Efron, tableau 6). Cependant, l'ampleur des biais issus des analyses CTN avec l'approximation d'Efron ayant diminué pour HR=1,25 par rapport à HR=2, l'erreur quadratique moyenne (RMSE) était plus favorable pour les estimations CTN avant qu'après réduction des biais pour tous les scénarios (tableau 6).

Tableau 6 – *Résultats des analyses CTN (pour 1 et 5 témoins par cas) avec réduction des biais selon la méthode de Firth : exposition fixe dans le temps et hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-7,078	0,269	0,267	95,7	-12,308	0,173	0,169	95,8	-21,038	0,102	0,111	93,2
CTN (1:1) RB	4,155	0,288	0,296	95,6	4,345	0,199	0,200	94,5	3,862	0,129	0,135	94,1
CTN (5:1) Ef	-5,243	0,215	0,225	94,0	-2,251	0,145	0,144	95,6	-2,824	0,091	0,091	95,2
CTN (5:1) RB	1,797	0,216	0,226	93,7	3,373	0,150	0,148	96,0	3,973	0,097	0,097	94,8
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-7,751	0,189	0,188	94,9	-11,707	0,123	0,126	94,3	-21,183	0,072	0,085	91,1
CTN (1:1) RB	1,664	0,203	0,207	94,6	4,073	0,141	0,148	94,4	2,198	0,091	0,094	94,2
CTN (5:1) Ef	-1,206	0,152	0,147	96,1	-2,046	0,104	0,105	95,0	-4,071	0,065	0,065	95,4
CTN (5:1) RB	1,948	0,154	0,149	95,7	1,745	0,107	0,108	94,9	1,666	0,069	0,068	95,1
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-6,928	0,169	0,165	94,7	-13,706	0,111	0,111	94,5	-21,156	0,065	0,080	88,3
CTN (1:1) RB	1,125	0,180	0,179	94,5	0,022	0,125	0,124	94,9	0,310	0,080	0,083	94,3
CTN (5:1) Ef	0,356	0,138	0,139	94,5	-1,175	0,094	0,092	95,2	-3,902	0,059	0,060	94,7
CTN (5:1) RB	1,598	0,140	0,141	94,7	1,621	0,096	0,094	95,2	1,145	0,062	0,063	95,3

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture; CTN cas-témoins niché; Ef Efron; RB réduction de biais; les résultats des lignes grisées sont les mêmes que dans le tableau 2 et sont inclus ici pour faciliter la comparaison.

- **Exposition variable dans le temps**

Des résultats similaires à ceux d'une exposition fixe dans le temps ont été obtenus pour une exposition variable dans le temps avec HR=2 (tableau 7) ou HR=1,25 (tableau 8).

Tableau 7 – *Résultats des analyses CTN (pour 1 et 5 témoins par cas) avec réduction des biais selon la méthode de Firth : exposition variable dans le temps et hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-12,283	0,335	0,328	95,2	-19,010	0,213	0,232	92,4	-27,395	0,125	0,218	70,6
CTN (1:1) RB	-0,315	0,373	0,383	95,4	1,675	0,263	0,268	96,0	2,552	0,175	0,177	95,4
CTN (5:1) Ef	-4,591	0,256	0,264	94,1	-4,080	0,174	0,165	96,5	-6,716	0,111	0,116	95,3
CTN (5:1) RB	-0,429	0,259	0,267	93,7	0,915	0,182	0,171	95,6	0,820	0,121	0,117	96,3
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-12,551	0,234	0,251	93,0	-17,568	0,150	0,186	88,3	-26,514	0,088	0,200	45,9
CTN (1:1) RB	-0,569	0,262	0,289	93,4	2,500	0,184	0,197	94,3	2,622	0,122	0,130	94,3
CTN (5:1) Ef	-3,458	0,182	0,191	94,3	-3,471	0,124	0,125	94,9	-7,119	0,079	0,091	90,8
CTN (5:1) RB	-0,668	0,185	0,194	94,2	0,730	0,129	0,130	94,3	-0,261	0,085	0,084	94,3
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-11,772	0,198	0,211	92,6	-17,178	0,128	0,171	85,7	-24,454	0,075	0,184	37,3
CTN (1:1) RB	-1,052	0,218	0,230	94,2	1,186	0,153	0,164	93,9	3,279	0,100	0,114	91,8
CTN (5:1) Ef	-3,611	0,157	0,160	93,8	-4,305	0,107	0,111	94,1	-5,983	0,067	0,081	89,2
CTN (5:1) RB	-1,629	0,160	0,162	93,8	-0,765	0,111	0,113	95,0	0,395	0,072	0,076	94,3

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture; CTN cas-témoins niché; Ef Efron; RB réduction de biais; les résultats des lignes grisées sont les mêmes que dans le tableau 3 et sont inclus ici pour faciliter la comparaison.

Tableau 8 – *Résultats des analyses CTN (pour 1 et 5 témoins par cas) avec réduction des biais selon la méthode de Firth : exposition variable dans le temps et hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-6,424	0,376	0,382	95,6	-15,887	0,242	0,234	96,4	-23,160	0,144	0,145	95,4
CTN (1:1) RB	5,372	0,396	0,417	95,1	1,047	0,276	0,278	95,8	1,551	0,182	0,181	94,9
CTN (5:1) Ef	-4,907	0,299	0,304	95,3	-2,774	0,203	0,195	96,5	-5,576	0,129	0,131	94,7
CTN (5:1) RB	8,012	0,298	0,302	94,9	5,652	0,207	0,200	96,3	2,238	0,136	0,138	94,7
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-13,185	0,252	0,250	95,1	-16,726	0,164	0,164	94,9	-23,891	0,098	0,110	92,0
CTN (1:1) RB	-3,690	0,269	0,273	95,2	-1,908	0,187	0,189	94,9	-1,126	0,123	0,128	93,8
CTN (5:1) Ef	-7,342	0,202	0,206	95,4	-3,718	0,138	0,140	95,7	-9,278	0,088	0,091	94,6
CTN (5:1) RB	-1,905	0,204	0,207	95,0	1,111	0,141	0,144	95,7	-3,356	0,093	0,095	95,1
50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
CTN (1:1) Ef	-12,709	0,203	0,208	94,3	-15,293	0,132	0,140	93,7	-22,324	0,079	0,094	91,5
CTN (1:1) RB	-4,063	0,217	0,228	94,4	-1,096	0,151	0,159	93,2	0,077	0,098	0,104	94,2
CTN (5:1) Ef	-9,748	0,163	0,171	94,4	-5,197	0,112	0,112	94,5	-6,585	0,071	0,075	94,8
CTN (5:1) RB	-7,118	0,165	0,173	94,3	-1,753	0,115	0,115	94,8	-1,192	0,075	0,078	94,4

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché ; Ef Efron ; RB réduction de biais ; les résultats des lignes grisées sont les mêmes que dans le tableau 4 et sont inclus ici pour faciliter la comparaison.

Dans tous les scénarios, les biais négatifs des estimations du log-HR issues des analyses CTN ont été considérablement réduits après l'utilisation de la méthode de réduction de biais, passant de 27% à 3% (tableau 7) pour un biais relatif maximal.

Dans tous les scénarios, la précision des estimations issues de l'analyse CTN avec la méthode de réduction de biais était à nouveau systématiquement inférieure à celle des estimations de l'analyse CTN avec l'approximation d'Efron. Les RMSE n'étaient ainsi avantageuses pour la méthode de réduction de biais que pour une plus grande proportion d'évènements (25%, tableau 7).

Les probabilités de couverture du schéma d'étude CTN qui étaient précédemment inférieures à la valeur nominale avec l'approximation d'Efron ont considérablement augmenté vers la valeur nominale après réduction des biais (tableau 7).

Le tableau 8 présentant les résultats des simulations pour un HR=1,25 montre une diminution considérable des biais dans les estimations des analyses CTN avec la méthode de réduction des biais (biais maximal étant de 5% alors qu'il était de 24% avec l'approximation d'Efron, tableau 8). Cependant, la légère diminution de l'ampleur des biais issus des analyses CTN avec l'approximation d'Efron pour HR=1,25 (par rapport à HR=2) a rendu l'erreur quadratique moyenne (RMSE) plus favorable pour les estimations CTN avec cette méthode d'approximation qu'avec la méthode de réduction de biais pour tous les scénarios (tableau 8).

4.1.3 Prise en compte des ex-æquos dans les analyses de la cohorte entière et cas-témoins nichées

- **Exposition fixe dans le temps**

Pour une exposition fixe dans le temps et un HR=2, la figure 3 représente le biais relatif des estimations du log-HR obtenues à partir des analyses CTN avec les approximations de Breslow et d'Efron et la méthode exacte pour la gestion des temps d'évènement ex-æquos dans les strates des échantillons CTN. Le tableau 9 complète ces résultats avec la moyenne des écarts-types estimés, l'erreur quadratique moyenne et la probabilité de couverture.

Tableau 9 – *Résultats des analyses de la cohorte entière (avec l'approximation d'Efron) et CTN (pour 1 et 5 témoins par cas) avec la méthode exacte, les approximations de Breslow et d'Efron et l'approche ccwc modifiée) pour une exposition fixe dans le temps et un hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-0,641	0,167	0,166	95,3	0,190	0,116	0,113	95,5	0,379	0,076	0,075	95,5
Cohorte Ef	-0,636	0,167	0,166	95,3	0,200	0,116	0,113	95,5	0,406	0,076	0,076	95,4
CTN (1:1) B	-16,227	0,242	0,249	93,1	-28,039	0,153	0,231	80,5	-41,119	0,090	0,292	4,6
CTN (1:1) Ef	-10,256	0,244	0,251	93,8	-18,677	0,155	0,194	87,6	-27,263	0,091	0,205	45,7
CTN (1:1) Ex	3,326	0,278	0,293	94,7	1,309	0,191	0,194	95,2	0,878	0,125	0,124	94,6
CTN (1:1) *	3,059	0,279	0,284	95,3	1,105	0,192	0,188	96,2	1,135	0,127	0,129	94,0
CTN (5:1) B	-4,120	0,188	0,181	95,1	-6,987	0,127	0,129	95,5	-11,896	0,081	0,111	84,2
CTN (5:1) Ef	-2,311	0,188	0,184	95,0	-3,673	0,128	0,128	95,4	-6,190	0,081	0,092	91,6
CTN (5:1) Ex	-0,022	0,193	0,189	95,1	0,274	0,134	0,131	95,3	0,610	0,088	0,089	95,4
CTN (5:1) *	-0,591	0,193	0,188	95,9	0,614	0,134	0,126	96,0	1,062	0,088	0,087	95,2
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-0,502	0,130	0,129	95,4	-0,085	0,090	0,089	95,8	0,113	0,058	0,057	96,0
Cohorte Ef	-0,498	0,130	0,129	95,4	-0,076	0,090	0,089	95,8	0,137	0,058	0,057	96,0
CTN (1:1) B	-15,659	0,181	0,195	93,2	-24,881	0,116	0,197	72,8	-38,109	0,068	0,269	0,2
CTN (1:1) Ef	-10,193	0,182	0,187	94,0	-15,829	0,117	0,154	86,1	-24,480	0,068	0,181	28,4
CTN (1:1) Ex	0,481	0,202	0,201	95,3	1,209	0,139	0,136	95,9	0,550	0,090	0,088	94,8
CTN (1:1) *	1,640	0,203	0,200	96,1	0,731	0,140	0,139	95,1	0,887	0,091	0,091	95,7
CTN (5:1) B	-3,948	0,143	0,144	95,4	-6,259	0,098	0,102	94,0	-10,972	0,061	0,095	77,8
CTN (5:1) Ef	-2,361	0,143	0,145	95,4	-3,328	0,098	0,098	95,0	-5,812	0,061	0,073	91,0
CTN (5:1) Ex	-0,472	0,146	0,147	94,9	0,046	0,101	0,099	96,1	0,100	0,066	0,065	94,3
CTN (5:1) *	-0,007	0,146	0,146	95,5	0,319	0,101	0,101	95,2	0,188	0,066	0,065	96,2

Tableau 9, suite

50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	0,204	0,135	0,137	94,8	0,228	0,093	0,091	94,8	0,069	0,059	0,059	95,6
Cohorte Ef	0,207	0,135	0,137	94,8	0,235	0,093	0,091	94,8	0,089	0,059	0,059	95,5
CTN (1:1) B	-13,234	0,177	0,192	92,6	-21,738	0,115	0,180	76,0	-33,994	0,067	0,242	2,8
CTN (1:1) Ef	-8,408	0,177	0,187	93,9	-13,388	0,116	0,143	88,8	-20,890	0,068	0,158	42,1
CTN (1:1) Ex	-0,345	0,190	0,194	94,9	-0,110	0,130	0,126	95,9	-0,227	0,082	0,081	95,1
CTN (1:1) *	0,557	0,191	0,198	93,9	-0,523	0,131	0,130	95,6	0,277	0,084	0,083	95,8
CTN (5:1) B	-2,741	0,144	0,146	94,9	-4,922	0,098	0,099	94,5	-9,192	0,061	0,087	82,5
CTN (5:1) Ef	-1,397	0,144	0,147	94,5	-2,446	0,098	0,097	95,4	-4,752	0,061	0,070	91,5
CTN (5:1) Ex	0,093	0,146	0,149	94,0	0,292	0,101	0,098	96,0	0,119	0,064	0,065	94,9
CTN (5:1) *	0,342	0,146	0,147	94,9	0,221	0,101	0,100	95,3	-0,046	0,064	0,065	94,7

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture; CTN cas-témoins niché ; B Breslow; Ef Efron ; Ex Exact ; * approche *ccwc* modifiée ; les résultats des lignes grisées sont les mêmes que dans le tableau 1 et sont inclus ici pour faciliter la comparaison.

Nous avons constaté que la fréquence des temps d'évènement ex-æquos était corrélée à la proportion d'évènements non censurés, allant d'une valeur médiane de 15% à 53% respectivement pour 5% à

25% de sujets qui ont observé un évènement, quelle que soit la prévalence de l'exposition (voir tableau 10).

Tableau 10 – *Distribution des temps d'évènements pour 1000 ensembles de données simulées en fonction des proportions d'évènements pour une exposition fixe dans le temps et un hazard ratio théorique de 2*

	5% d'évènements			10% d'évènements			25% d'évènements		
	Médiane	Min	Max	Médiane	Min	Max	Médiane	Min	Max
Nombre total d'évènements, M	249	202	313	521	456	593	1,265	1,173	1,363
Nombre de temps d'évènement distincts, K	211	174	259	373	331	413	600	567	636
Nombre de temps d'évènement uniques, M_1	177	149	213	255	211	296	223	178	263
Proportion des temps d'évènement ex-æquos, $(M-K)/M$	15%	8%	23%	28%	23%	33%	53%	50%	56%

Les résultats étaient identiques pour les trois proportions de sujets exposés considérées

Dans tous les scénarios, les estimations du log-HR étaient sans biais (biais relatif <1 %) pour l'analyse de la cohorte entière avec les approximations d'Efron ou de Breslow (tableau 9). En revanche, les estimations du log-HR pour les analyses CTN présentaient un biais négatif systématique, souvent important, avec l'approximation d'Efron et encore plus avec l'approximation de Breslow (figure 3A, B et C, D respectivement). Ce biais était beaucoup plus prononcé avec l'appariement 1 :1 qu'avec l'appariement 5:1 (par exemple, le biais relatif maximal était de 41% pour 1 témoin par cas contre 12% pour 5 témoins par cas avec l'approximation de Breslow, figure 3A et B respectivement). Ce biais augmentait sensiblement avec la proportion d'évènements (par exemple, pour l'approximation de Breslow et 25% de sujets exposés, le biais relatif était de 15 %, 25% et 38% respectivement pour 5 %, 10% et 25% de sujets ayant subi un évènement, figure 3A). De plus, le biais diminuait légèrement avec l'augmentation de la proportion de sujets exposés (par exemple, pour l'approximation d'Efron et 25% de sujets ayant subi un évènement, le biais relatif était de 27 %, 24% et 21% pour respectivement 10%, 25% et 50% de sujets exposés, figure 3C). Ces biais étaient presque entièrement éliminés lorsqu'on utilisait la méthode exacte pour la gestion des ex-æquos ou l'approche ccwc modifiée. En effet, le biais relatif maximal était seulement de 3% pour les deux méthodes (figure 3E, tableau 9) contre 41% et 27% pour les approximations de Breslow et d'Efron, respectivement (figure 3A et C respectivement, tableau 9).

Dans tous les scénarios, les estimations basées sur l'analyse de la cohorte entière ont présenté une variance systématiquement plus faible que celles des analyses CTN (voir les colonnes SEE des écarts-types moyens estimés dans le tableau 9). Comme prévu, la précision des estimations s'est améliorée de façon générale pour les analyses des deux schémas d'étude à mesure que la proportion d'évènements, la prévalence de l'exposition et (pour les analyses CTN) le nombre de témoins par cas augmentaient. La précision des estimations issues de l'analyse CTN avec la méthode exacte ou l'approche *ccwc* modifiée était systématiquement inférieure à celle des estimations de l'analyse CTN

avec l'approximation de Breslow ou d'Efron. Par conséquent, le compromis biais-variance n'était en faveur de la méthode exacte ou de l'approche *ccwc* modifiée que pour une plus grande proportion d'évènements, c'est-à-dire 25% et peut-être 10% (voir les colonnes RMSE dans le tableau 9).

Dans tous les scénarios, nous avons observé des niveaux nominaux (0,95) pour la probabilité de couverture du schéma de cohorte tandis que le schéma d'étude CTN présentait des probabilités de couverture de niveaux inférieurs à la valeur nominale, en particulier pour les approximations de Breslow et d'Efron avec des proportions d'évènements plus élevées (scénarios pour lesquels des biais considérables ont été observés).

Les résultats de la simulation pour HR=1,25 étaient globalement similaires à ceux pour HR=2. Les estimations du log-HR pour l'analyse de la cohorte entière (avec les approximations de Breslow ou d'Efron) étaient non biaisées sauf pour le scénario avec 10%de sujets exposés et 5%d'évènements où on a observé un léger biais négatif de 5% (tableau 11). Bien que leur ampleur soit quelque peu réduite, un biais négatif augmentant avec une proportion plus élevée d'évènements a de nouveau été observé dans les analyses CTN utilisant l'approximation d'Efron ou de Breslow. Cependant, comme leur biais était plus faible, le compromis biais-variance est devenu plus favorable et les probabilités de couverture étaient plus proches de la valeur nominale pour ces deux méthodes d'approximation avec un HR=1,25 (tableau 11).

Tableau 11 – *Résultats des analyses de la cohorte entière (avec l'approximation d'Efron) et CTN (pour 1 et 5 témoins par cas) avec la méthode exacte, les approximations de Breslow et d'Efron et l'approche ccwc modifiée) pour une exposition fixe dans le temps et un hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-5,335	0,196	0,204	94,6	0,520	0,135	0,136	94,3	1,674	0,087	0,087	94,7
Cohorte Ef	-5,331	0,196	0,204	94,6	0,529	0,135	0,136	94,3	1,697	0,087	0,087	94,7
CTN (1:1) B	-12,671	0,267	0,251	96,3	-22,057	0,172	0,155	96,9	-35,066	0,101	0,113	92,7
CTN (1:1) Ef	-7,078	0,269	0,267	95,7	-12,308	0,173	0,169	95,8	-21,038	0,102	0,111	93,2
CTN (1:1) Ex	5,851	0,293	0,300	95,7	5,051	0,201	0,201	94,4	4,004	0,129	0,135	94,3
CTN (1:1) *	4,955	0,293	0,300	95,2	2,498	0,201	0,208	94,5	2,194	0,129	0,129	96,4
CTN (5:1) B	-6,708	0,215	0,222	94,3	-5,176	0,145	0,140	95,9	-7,890	0,091	0,087	96,0
CTN (5:1) Ef	-5,243	0,215	0,225	94,0	-2,251	0,145	0,144	95,6	-2,824	0,091	0,091	95,2
CTN (5:1) Ex	-3,206	0,218	0,229	93,8	2,043	0,150	0,149	95,7	2,970	0,097	0,097	94,9
CTN (5:1) *	-4,253	0,218	0,225	94,5	2,392	0,150	0,150	95,7	1,606	0,097	0,096	95,4
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-0,868	0,140	0,139	94,8	0,158	0,097	0,099	94,3	0,541	0,062	0,062	96,4
Cohorte Ef	-0,865	0,140	0,139	94,8	0,166	0,097	0,099	94,3	0,562	0,062	0,062	96,4
CTN (1:1) B	-13,105	0,188	0,178	95,9	-20,948	0,122	0,120	95,8	-34,881	0,072	0,097	88,0
CTN (1:1) Ef	-7,751	0,189	0,188	94,9	-11,707	0,123	0,126	94,3	-21,183	0,072	0,085	91,1
CTN (1:1) Ex	2,036	0,204	0,207	94,6	3,894	0,141	0,147	94,7	1,832	0,090	0,093	94,3
CTN (1:1) *	1,440	0,204	0,200	95,6	2,763	0,141	0,143	94,6	3,986	0,091	0,091	95,1
CTN (5:1) B	-2,674	0,152	0,145	96,6	-4,796	0,103	0,102	95,3	-8,862	0,065	0,064	95,7
CTN (5:1) Ef	-1,206	0,152	0,147	96,1	-2,046	0,104	0,105	95,0	-4,071	0,065	0,065	95,4
CTN (5:1) Ex	0,549	0,155	0,150	95,8	1,025	0,107	0,108	95,0	1,345	0,069	0,068	95,1
CTN (5:1) *	-0,772	0,155	0,155	94,9	0,858	0,107	0,108	94,8	1,208	0,069	0,067	95,4

Tableau 11, suite

50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	1,917	0,128	0,128	95,4	1,787	0,088	0,087	95,0	0,559	0,057	0,057	95,1
Cohorte Ef	1,921	0,128	0,128	95,4	1,795	0,088	0,087	95,0	0,579	0,057	0,057	95,0
CTN (1:1) B	-12,072	0,169	0,157	96,0	-22,283	0,110	0,108	95,1	-34,468	0,065	0,094	81,8
CTN (1:1) Ef	-6,928	0,169	0,165	94,7	-13,706	0,111	0,111	94,5	-21,156	0,065	0,080	88,3
CTN (1:1) Ex	1,358	0,181	0,179	94,7	-0,199	0,125	0,123	94,9	-0,108	0,080	0,082	94,7
CTN (1:1) *	-0,782	0,182	0,179	94,9	1,389	0,125	0,122	95,4	1,597	0,080	0,081	94,4
CTN (5:1) B	-1,025	0,138	0,137	94,9	-3,720	0,094	0,090	95,7	-8,466	0,059	0,060	94,3
CTN (5:1) Ef	0,356	0,138	0,139	94,5	-1,175	0,094	0,092	95,2	-3,902	0,059	0,060	94,7
CTN (5:1) Ex	1,972	0,140	0,141	94,7	1,725	0,097	0,094	95,2	1,130	0,062	0,062	95,3
CTN (5:1) *	2,313	0,140	0,140	95,2	1,873	0,097	0,094	94,7	0,371	0,062	0,062	96,6

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture; CTN cas-témoins niché ; B Breslow ; Ef Efron ; Ex Exact ; * approche *ccwc* modifiée ; les résultats des lignes grisées sont les mêmes que dans le tableau 2 et sont inclus ici pour faciliter la comparaison.

- **Exposition variable dans le temps**

Les résultats pour une exposition variable dans le temps étaient similaires à ceux obtenus pour une exposition fixe, que ce soit avec un HR=2 (tableau 12) ou un HR=1,25 (tableau 13). La figure 4 représente les biais relatifs des estimations du log-HR obtenues à partir des analyses CTN avec les approximations de Breslow et d'Efron et la méthode exacte.

Tableau 12 – *Résultats des analyses de la cohorte entière (avec l'approximation d'Efron) et CTN (pour 1 et 5 témoins par cas) avec la méthode exacte, les approximations de Breslow et d'Efron et l'approche ccwc modifiée) pour une exposition variable dans le temps et un hazard ratio théorique de 2*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-2,333	0,226	0,226	96,1	-0,621	0,157	0,146	96,9	0,478	0,104	0,100	96,8
Cohorte Ef	-2,328	0,226	0,226	96,1	-0,611	0,157	0,146	96,9	0,506	0,104	0,100	96,7
CTN (1:1) B	-18,475	0,331	0,318	95,9	-28,404	0,210	0,258	89,5	-41,108	0,123	0,297	29,2
CTN (1:1) Ef	-12,283	0,335	0,328	95,2	-19,010	0,213	0,232	92,4	-27,395	0,125	0,218	70,6
CTN (1:1) Ex	2,200	0,385	0,390	96,0	2,265	0,265	0,264	96,0	1,689	0,174	0,168	96,0
CTN (1:1) *	2,188	0,385	0,383	96,1	2,061	0,266	0,265	95,5	1,602	0,175	0,182	94,7
CTN (5:1) B	-6,503	0,256	0,259	94,7	-7,464	0,174	0,164	96,7	-12,415	0,111	0,130	92,0
CTN (5:1) Ef	-4,591	0,256	0,264	94,1	-4,080	0,174	0,165	96,5	-6,716	0,111	0,116	95,3
CTN (5:1) Ex	-2,263	0,263	0,271	93,7	-0,003	0,183	0,172	96,1	0,127	0,121	0,116	96,5
CTN (5:1) *	-1,924	0,263	0,259	95,8	-0,526	0,183	0,173	96,8	0,481	0,121	0,123	95,7
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-1,620	0,163	0,170	94,2	0,173	0,113	0,113	95,1	-0,845	0,074	0,074	94,7
Cohorte Ef	-1,616	0,163	0,170	94,2	0,183	0,113	0,113	95,1	-0,820	0,074	0,074	94,5
CTN (1:1) B	-18,136	0,232	0,252	92,4	-26,770	0,148	0,223	79,6	-39,979	0,087	0,284	4,0
CTN (1:1) Ef	-12,551	0,234	0,251	93,0	-17,568	0,150	0,186	88,3	-26,514	0,088	0,200	45,9
CTN (1:1) Ex	-0,552	0,263	0,279	93,8	1,085	0,182	0,185	94,9	-0,093	0,119	0,116	95,6
CTN (1:1) *	0,992	0,265	0,273	94,2	2,183	0,183	0,187	95,4	-0,482	0,120	0,119	94,7
CTN (5:1) B	-5,158	0,182	0,189	94,3	-6,632	0,123	0,127	94,0	-12,451	0,078	0,112	84,5
CTN (5:1) Ef	-3,458	0,182	0,191	94,3	-3,471	0,124	0,125	94,9	-7,119	0,079	0,091	90,8
CTN (5:1) Ex	-1,344	0,186	0,195	94,3	0,226	0,129	0,129	94,5	-0,887	0,085	0,083	94,4
CTN (5:1) *	-1,313	0,186	0,192	94,3	0,074	0,129	0,132	94,6	-1,081	0,085	0,087	94,1

Tableau 12, suite

50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-1,770	0,143	0,147	94,0	-0,947	0,099	0,099	94,3	-0,067	0,064	0,066	94,6
Cohorte Ef	-1,766	0,143	0,147	94,0	-0,938	0,099	0,099	94,3	-0,044	0,064	0,066	94,6
CTN (1:1) B	-17,164	0,196	0,217	91,7	-25,875	0,127	0,210	72,6	-37,784	0,074	0,268	2,3
CTN (1:1) Ef	-11,772	0,198	0,211	92,6	-17,178	0,128	0,171	85,7	-24,454	0,075	0,184	37,3
CTN (1:1) Ex	-1,807	0,218	0,221	94,7	-1,127	0,150	0,151	94,2	-0,543	0,097	0,099	94,4
CTN (1:1) *	-0,228	0,220	0,220	95,7	-0,444	0,152	0,151	94,8	-0,233	0,098	0,097	95,7
CTN (5:1) B	-5,204	0,157	0,159	93,9	-7,162	0,107	0,115	92,9	-10,977	0,067	0,100	79,4
CTN (5:1) Ef	-3,611	0,157	0,160	93,8	-4,305	0,107	0,111	94,1	-5,983	0,067	0,081	89,2
CTN (5:1) Ex	-1,843	0,160	0,162	93,7	-1,079	0,111	0,112	95,0	-0,287	0,072	0,074	94,4
CTN (5:1) *	-2,668	0,160	0,165	94,2	-1,209	0,111	0,108	95,8	-0,235	0,072	0,074	94,6

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché; B Breslow; Ef Efron ; Ex Exact ; * approche *ccwc* modifiée ; les résultats des lignes grisées sont les mêmes que dans le tableau 3 et sont inclus ici pour faciliter la comparaison.

Tableau 13 – *Résultats des analyses de la cohorte entière (avec l'approximation d'Efron) et CTN (pour 1 et 5 témoins par cas) avec la méthode exacte, les approximations de Breslow et d'Efron et l'approche ccwc modifiée) pour une exposition variable dans le temps et un hazard ratio théorique de 1,25*

10% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-8,617	0,273	0,286	95,2	-2,728	0,188	0,177	97,1	-0,492	0,123	0,123	95,3
Cohorte Ef	-8,613	0,273	0,287	95,2	-2,719	0,188	0,177	97,1	-0,470	0,123	0,123	95,3
CTN (1:1) B	-12,790	0,374	0,358	96,3	-25,603	0,240	0,213	97,6	-36,669	0,143	0,138	96,8
CTN (1:1) Ef	-6,424	0,376	0,382	95,6	-15,887	0,242	0,234	96,4	-23,160	0,144	0,145	95,4
CTN (1:1) Ex	9,574	0,413	0,436	94,9	2,756	0,281	0,282	95,8	2,068	0,183	0,182	95,1
CTN (1:1) *	8,916	0,415	0,447	94,9	11,889	0,283	0,278	96,4	1,857	0,183	0,188	95,2
CTN (5:1) B	-6,540	0,299	0,299	95,5	-5,828	0,203	0,189	97,1	-10,494	0,129	0,125	95,8
CTN (5:1) Ef	-4,907	0,299	0,304	95,3	-2,774	0,203	0,195	96,5	-5,576	0,129	0,131	94,7
CTN (5:1) Ex	-2,696	0,305	0,310	95,1	0,554	0,210	0,202	96,3	0,082	0,137	0,138	94,6
CTN (5:1) *	-5,320	0,304	0,316	96,1	-3,578	0,209	0,199	96,4	-1,479	0,137	0,138	95,2
25% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-5,924	0,185	0,184	95,7	-1,131	0,128	0,129	95,5	-2,552	0,084	0,085	94,4
Cohorte Ef	-5,921	0,185	0,184	95,7	-1,122	0,128	0,129	95,5	-2,531	0,084	0,085	94,3
CTN (1:1) B	-18,547	0,250	0,235	96,4	-25,853	0,163	0,152	96,3	-37,302	0,097	0,115	91,7
CTN (1:1) Ef	-13,185	0,252	0,250	95,1	-16,726	0,164	0,164	94,9	-23,891	0,098	0,110	92,0
CTN (1:1) Ex	-2,509	0,273	0,276	95,0	-1,371	0,189	0,189	95,1	-1,387	0,123	0,127	94,0
CTN (1:1) *	0,331	0,274	0,287	94,1	1,535	0,189	0,189	95,2	-0,792	0,123	0,125	94,9
CTN (5:1) B	-8,783	0,202	0,202	95,6	-6,492	0,138	0,136	96,1	-13,880	0,088	0,090	94,9
CTN (5:1) Ef	-7,342	0,202	0,206	95,4	-3,718	0,138	0,140	95,7	-9,278	0,088	0,091	94,6
CTN (5:1) Ex	-5,558	0,205	0,209	95,2	-0,586	0,142	0,144	95,6	-4,185	0,093	0,095	94,8
CTN (5:1) *	-5,777	0,205	0,201	95,3	-0,584	0,142	0,141	95,9	-3,252	0,093	0,094	94,2

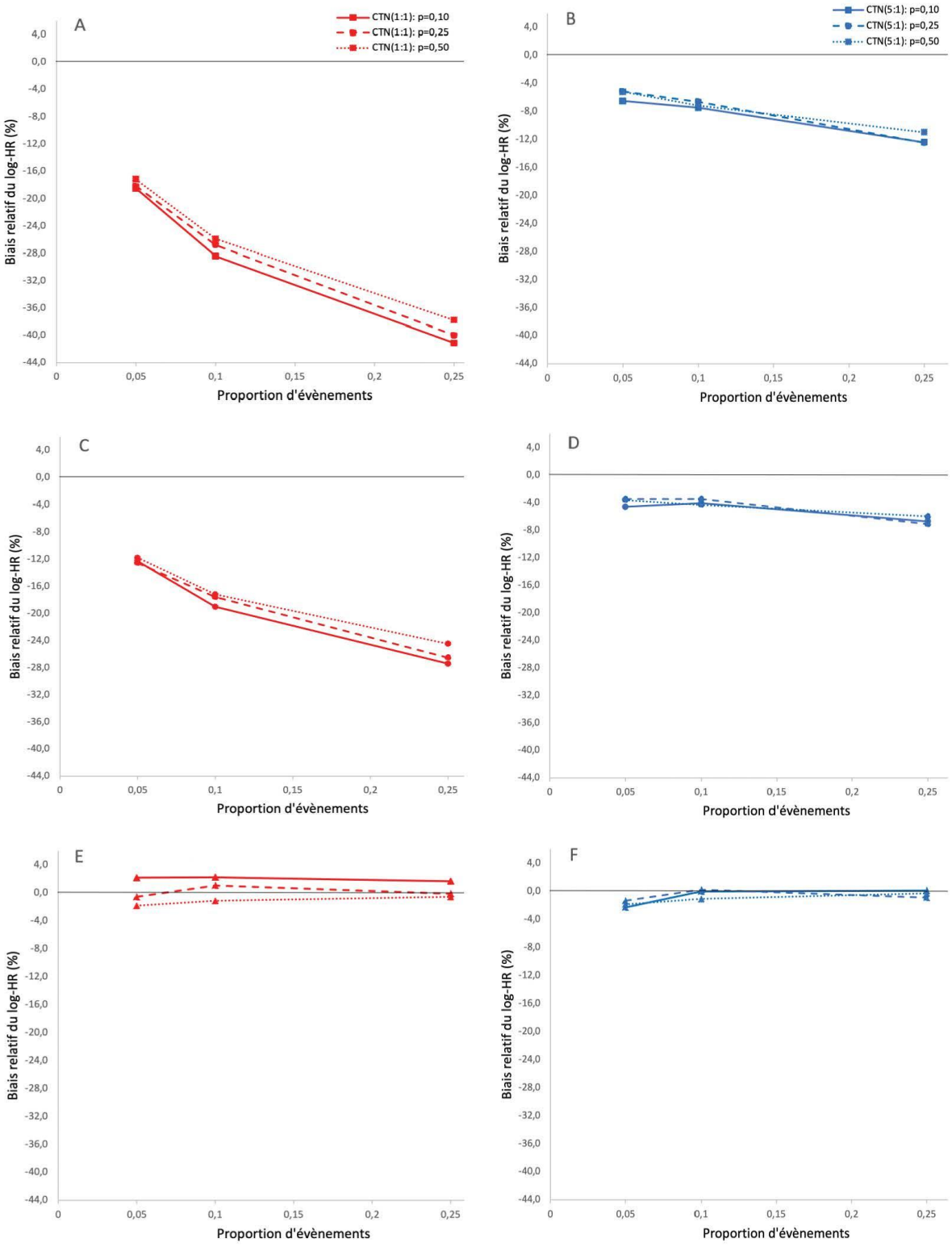
Tableau 13, suite

50% sujets exposés												
Schéma d'étude	5% d'évènements				10% d'évènements				25% d'évènements			
	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)	Biais relatif (%)	SEE	RMSE	PC (%)
Cohorte B	-6,147	0,151	0,158	94,7	-1,645	0,104	0,105	94,8	-0,380	0,068	0,071	93,9
Cohorte Ef	-6,144	0,151	0,158	94,7	-1,637	0,104	0,105	94,8	-0,359	0,068	0,071	93,9
CTN (1:1) B	-17,945	0,202	0,198	95,3	-24,273	0,132	0,132	94,5	-35,610	0,078	0,103	87,1
CTN (1:1) Ef	-12,709	0,203	0,208	94,3	-15,293	0,132	0,140	93,7	-22,324	0,079	0,094	91,5
CTN (1:1) Ex	-3,622	0,219	0,228	94,3	-0,983	0,151	0,159	93,4	-0,405	0,098	0,103	94,4
CTN (1:1) *	-5,338	0,219	0,222	95,5	-1,120	0,151	0,155	94,9	0,493	0,098	0,100	94,7
CTN (5:1) B	-11,170	0,163	0,169	94,5	-7,818	0,112	0,110	94,9	-11,170	0,071	0,074	94,0
CTN (5:1) Ef	-9,748	0,163	0,171	94,4	-5,197	0,112	0,112	94,5	-6,585	0,071	0,075	94,8
CTN (5:1) Ex	-8,245	0,166	0,174	94,2	-2,284	0,115	0,115	94,8	-1,519	0,075	0,078	94,7
CTN (5:1) *	-6,293	0,166	0,177	92,9	-1,318	0,115	0,115	95,7	-0,511	0,075	0,077	94,7

SEE *average standard error estimator* ; RMSE *root mean squared error* ; PC probabilité de couverture ; CTN cas-témoins niché ; B Breslow ;
Ef Efron ; Ex Exact ; * approche *cowc* modifiée ; les résultats des lignes grisées sont les mêmes que dans le tableau 4 et sont inclus ici pour faciliter la comparaison.

Biopark - 8, rue de la croix Jarry - 75013 Paris
info@iresp.net - www.iresp.net

Figure 4 – *Biais relatif du log-HR des analyses CTN (pour 1 et 5 témoins par cas) avec les approximations de Breslow (A et B respectivement) et d'Efron (C et D respectivement) et la méthode exacte (E et F respectivement) : exposition variable dans le temps et hazard ratio théorique de 2*



La fréquence des temps d'évènement ex-æquos était comparable à celle observée avec une exposition fixe dans les neuf scénarios pour HR=2 (tableau 14) considérés.

Tableau 14 – *Distribution des temps d'évènements pour 1000 ensembles de données simulées en fonction des proportions d'évènements et de sujets exposés pour une exposition variable dans le temps et un hazard ratio théorique de 2*

	5% d'évènements			10% d'évènements			25% d'évènements		
	Médiane	Min	Max	Médiane	Min	Max	Médiane	Min	Max
10% sujets exposés									
Nombre total d'évènements, M	249	202	314	521	462	592	1,265	1,172	1,354
Nombre de temps d'évènement distincts, K	211	174	259	373	334	420	600	569	630
Nombre de temps d'évènement uniques, M_1	178	147	212	255	218	298	223	182	259
Proportion des temps d'évènement ex-æquos, $(M\text{-}K)/M$	15%	8%	22%	28%	23%	33%	53%	49%	56%
25% sujets exposés									
	5% d'évènements			10% d'évènements			25% d'évènements		
	Médiane	Min	Max	Médiane	Min	Max	Médiane	Min	Max
Nombre total d'évènements, M	249	202	302	522	453	585	1,265	1,179	1,366
Nombre de temps d'évènement distincts, K	211	175	251	372	330	416	600	571	631
Nombre de temps d'évènement uniques, M_1	177	147	216	255	212	300	223	184	263
Proportion des temps d'évènement ex-æquos, $(M\text{-}K)/M$	15%	8%	23%	28%	22%	35%	53%	49%	56%
50% sujets exposés									
	5% d'évènements			10% d'évènements			25% d'évènements		
	Médiane	Min	Max	Médiane	Min	Max	Médiane	Min	Max
Nombre total d'évènements, M	249	204	300	522	451	582	1,266	1,171	1,367
Nombre de temps d'évènement distincts, K	211	174	248	372	332	407	600	566	628
Nombre de temps d'évènement uniques, M_1	177	137	213	255	215	293	223	184	264
Proportion des temps d'évènement ex-æquos, $(M\text{-}K)/M$	15%	8%	23%	29%	23%	35%	53%	50%	56%

L'analyse de la cohorte entière avec les approximations de Breslow et d'Efron a donné des estimations non biaisées (biais relatif <2,5 %, tableau 12), tandis que les estimations issues des analyses CTN avec les approximations de Breslow et d'Efron étaient à nouveau biaisées, parfois fortement (biais relatif maximal de 41% pour un témoin par cas et 12% pour 5 témoins par cas pour l'approximation de Breslow,

figure 4A et B respectivement). La même tendance d'un biais croissant à mesure que la proportion d'évènements augmente était apparente (par exemple, pour l'approximation d'Efron et pour 10% de sujets exposés, le biais est passé de 12% à 19% puis à 27% pour respectivement 5 %, 10% et 25% de sujets ayant subi un évènement, figure 4C). Il n'y avait cependant pas de tendance claire avec la proportion de sujets exposés. L'approximation de Breslow pour la gestion des temps d'évènement ex-æquos a entraîné un biais plus important dans les estimations du schéma CTN que l'approximation d'Efron dans tous les scénarios (figure 4A, B et C, D respectivement, tableau 12). Ces biais ont été considérablement réduits avec l'utilisation de la méthode exacte ou de l'approche *ccwc* modifiée avec un seul cas dans chaque strate (biais relatif maximal de 41% et 27% pour les approximations de Breslow et d'Efron respectivement à 2% pour la méthode exacte et l'approche *ccwc* modifiée, figure 4A, C et E respectivement, tableau 12).

Le tableau 12 montre des écarts-types moyens estimés (SEE) et des erreurs quadratiques moyennes (RMSE) plus faibles pour les estimations issues de l'analyse de la cohorte entière que pour celles de l'analyse CTN dans tous les scénarios. La précision des estimations s'est améliorée pour les deux schémas d'étude à mesure que la proportion d'évènements, la prévalence de l'exposition et, pour les analyses CTN, le nombre de témoins par cas augmentaient. Comme avec l'exposition fixe dans le temps, pour l'analyse CTN avec les approximations de Breslow et d'Efron où les estimations ont produit un biais plus important, les RMSE étaient plus grandes pour une proportion plus élevée d'évènements et un témoin par cas et les SEE étaient systématiquement plus petites que celles de l'analyse CTN avec la méthode exacte ou l'approche *ccwc* modifiée (tableau 12).

De même que ce que nous avons observé pour une exposition fixe, la probabilité de couverture des estimations de la cohorte était proche de 95% pour tous les scénarios, tandis que celle des estimations CTN était en dessous de la valeur nominale, en particulier pour les approximations de Breslow et d'Efron avec des proportions plus élevées d'évènements et de sujets exposés (scénarios pour lesquels les biais les plus importants ont été observés).

Les résultats des simulations avec un HR=1,25 révèlent que les estimations du log-HR pour l'analyse de la cohorte entière étaient non biaisées. Cependant, pour les scénarios avec 5% d'évènements, on a observé un biais négatif compris entre 5% et 9% (tableau 13). Nous avons également observé un biais négatif (légèrement plus faible que celui obtenu pour un HR=2) qui augmentait avec la proportion d'évènements (pour l'approximation de Breslow, 25% de sujets exposés et un témoin par cas, le biais relatif passait de 18% à 37% pour 5% et 25% d'évènements respectivement) dans les résultats des analyses CTN (tableau 13).

4.1.4 Discussion et perspectives

- **Synthèse des résultats**

L'objectif de la thèse de Liliane Manitchoko était d'évaluer les méthodes statistiques disponibles pour modéliser les expositions médicamenteuses et leurs effets variant dans le temps. Plus spécifiquement, il était d'examiner les performances des schémas d'étude de cohorte et cas-témoins niché pour estimer l'association entre une exposition médicamenteuse et le risque d'évènement indésirable, en particulier pour une exposition variable dans le temps.

Les résultats de l'étude de simulation que nous avons réalisée ont montré que les estimations obtenues de l'analyse de la cohorte entière étaient non biaisées dans tous les scénarios considérés. Cependant, les estimations issues des analyses CTN (avec l'approximation d'Efron) étaient substantiellement biaisées, en particulier lorsqu'un seul témoin était apparié à chaque cas. Ce biais des estimations CTN augmentait avec la proportion d'évènements.

Pour comprendre d'où venaient les biais dans les résultats des analyses CTN et y remédier éventuellement, nous avons dans un premier temps analysé les données CTN avec la méthode de Firth adaptée à la régression logistique conditionnelle pour la réduction des biais. Nous avons constaté une diminution considérable des biais dans les estimations suggérant ainsi que les biais observés pourraient provenir de données éparses.

Cependant, certains arguments nous ont fait comprendre que cette explication n'était pas satisfaisante. En effet, après avoir augmenté la taille des cohortes simulées (résultats non présentés dans ce rapport) et surtout quel que soit le nombre d'évènements, les biais persistaient alors qu'en principe plus la proportion d'évènements augmente plus le risque de données éparses diminue. Nous avons donc, dans un second temps, pris en compte les temps d'évènement exæquos (pouvant apparaître dans les strates) en utilisant différentes méthodes : les approximations de Breslow et d'Efron (cette dernière étant celle utilisée par défaut dans les premières analyses CTN et de la cohorte entière), la méthode exacte et l'approche *ccwc* modifiée. Nous avons constaté que les estimations CTN avec l'approximation de Breslow étaient davantage biaisées qu'avec l'approximation d'Efron, tandis qu'avec la méthode exacte ou l'approche *ccwc* modifiée les estimations étaient presque exemptes de biais et assez proches de celles de l'analyse de la cohorte entière. Il est ainsi apparu que les biais initialement observés résultaient vraisemblablement de la présence de temps d'évènement ex-æquos dans les strates. Ces ex-æquos sont présents dans la cohorte entière mais relativement "dilués" de sorte que leur prise en compte impacte peu l'estimation de l'effet de l'exposition. En revanche, avec la méthode de sélection des témoins utilisée pour l'extraction des données CTN, ces ex-æquos se retrouvaient dans des strates réduites aux cas et à leurs témoins appariés, de sorte que les approximations de Breslow et d'Efron se trouvaient très éloignées de la vraisemblance qu'elles étaient censées approcher.

Quel que soit le scénario considéré, les estimations basées sur l'analyse de la cohorte entière ont présenté une précision systématiquement plus faible que celles des analyses CTN. Les estimations issues de l'analyse de la cohorte entière avaient un taux de recouvrement satisfaisant tandis que celles issues de l'analyse CTN présentaient de faibles taux de recouvrement, en particulier avec les

approximations de Breslow et d'Efron pour des scénarios où des biais considérables avaient été observés. Tous ces résultats ont été observés autant pour une exposition variable que fixe dans le temps.

- **Comparaison des résultats à ceux de la littérature existante**

Les résultats très biaisés obtenus après les analyses CTN (1 :1) (avec l'approximation d'Efron ou de Breslow) des données simulées par rapport à ceux des analyses de la cohorte entière contredisent la majorité des résultats de la littérature pour lesquels les analyses de données selon ces deux schémas d'étude produisent des estimations presque semblables (Billioti de Gage *et al* 2012 ; Breslow et al 1983 ; Essebag *et al* 2005 ; Hak *et al* 2004 ; Liddell *et al* 1977). Cependant, de tels résultats ont été retrouvés dans les quelques travaux (à notre connaissance) qui comparaient ces deux schémas d'étude à l'aide des données simulées avec une exposition fixe ou variable dans le temps (Austin *et al* 2012 ; Bertke *et al* 2013 ; Leffondré *et al* 2003). L'utilisation de la méthode de Firth adaptée à la régression logistique conditionnelle pour la réduction des biais dans les analyses CTN a considérablement réduit ces biais et a ainsi montré qu'ils pouvaient être dus aux données éparses. Ces résultats sont similaires à ceux de l'étude de Cheung *et al* (2019) qui ont montré que des estimations biaisées obtenues dans les analyses CTN étaient dues au biais des données éparses qui pouvait être contrôlé par cette méthode de réduction de biais pour la régression logistique conditionnelle.

La perte de précision des estimations issues des analyses CTN par rapport à celles des analyses de la cohorte entière était attendue et corrobore avec les résultats de travaux antérieurs (Breslow *et al* 1983; Essebag *et a*l 2005 ; Hak *et a*l 2004 ; Liddell *et al* 1977) car, dans le modèle d'analyse de cohorte, toutes les données sont utilisées alors que, dans celui de l'analyse CTN, seules les données sur les cas et les témoins sélectionnés sont utilisées.

Le biais encore apparent avec 5 témoins par cas était cependant moins attendu et contradictoire avec l'idée générale selon laquelle les analyses CTN avec un nombre suffisant de témoins par cas fournissent des estimations proches de celles des analyses de cohortes entières (Essebag *et al* 2005 ; Bertke *et al* 2013 ; Goldstein *et al* 1992 ; Pang 1999). De plus, de manière contre-intuitive, le biais a eu tendance à augmenter avec une plus grande proportion d'évènements, un résultat similaire à celui rapporté par Austin *et al.* (2012) alors qu'on pourrait s'attendre à une estimation plus précise lorsque la richesse de l'information augmente comme observé dans les résultats de l'étude de Bertke et al (2013). Dans notre cas cependant, les biais observés résultaient très probablement d'évènements ex-æquos et de la façon dont ils étaient traités. Ces biais étaient en effet nettement réduits (au prix d'une précision légèrement inférieure pour une proportion plus faible d'évènements) lorsque la méthode exacte (Peto 1972) ou l'approche *ccwc* modifiée étaient utilisées au lieu de l'approximation de Breslow (1974) ou d'Efron (1977).

La tendance observée d'une augmentation des biais avec une plus grande proportion d'évènements avec les approximations de Breslow et d'Efron est cohérente avec les résultats de travaux antérieurs montrant que ces approximations se détériorent lorsque la proportion d'évènements ex-æquos

augmente (Farewell *et al* 1980). Dans notre cadre de simulation, la probabilité que deux évènements ou plus se produisent en même temps dans un délai limité de 730 jours augmente avec la proportion d'évènements, c'est-à-dire le nombre de cas dans une cohorte de taille fixe de 5 000 individus. L'augmentation de la proportion d'évènements ex-æquos a eu un fort impact négatif sur les estimations des analyses des données CTN, alors que les estimations de l'analyse de la cohorte entière ont été à peine affectées et sont restées pratiquement sans biais. En effet, dans les analyses de cohorte entière, la proportion d'ex-æquos dans les ensembles à risque à chaque temps d'évènement est considérablement diluée parmi tous les sujets appartenant à l'ensemble à risque à ce temps-là, alors que dans les analyses CTN, l'ensemble à risque est beaucoup plus petit car limité aux cas survenant à ce temps spécifique de l'évènement et à leurs témoins appariés. Auparavant, Farewell et Prentice (1980) avaient montré par des simulations de données cas-témoins avec 5, 20 et 50 cas (c'est-à-dire des évènements ex-æquos) par strate que les estimations des analyses CTN avec l'approximation d'Efron, et dans une plus large mesure encore avec l'approximation de Breslow, sont affectées par un biais de sous-estimation systématique. Ils ont donc conseillé de ne pas utiliser ces deux méthodes d'approximation pour les études cas-témoins ou de cohorte dans lesquelles le nombre d'ex-æquos par temps d'évènement est important. D'autres auteurs ont préconisé l'utilisation de la méthode exacte (Peto 1972) pour analyser les études cas-témoins lorsque la proportion de cas dans chaque strate cas-témoins n'est pas suffisamment faible (Langholz & Richardson 2010). Ce problème peut également être évité dans les études cas-témoins en disposant les ensembles cas-témoins de manière à obtenir un seul cas par strate et, ainsi, à ne pas avoir d'évènements ex-æquos dans les strates. Cette solution exige toutefois que chacun des cas survenant au même moment de l'évènement puisse se voir attribuer un nombre suffisant de témoins appariés. Cela a motivé notre modification de la fonction *ccwc* dans R, conformément à la procédure d'échantillonnage réalisée dans la macro SAS %nCCsampling (Desai *et al* 2016).

Nous craignons que les chercheurs ne soient pas conscients de la détérioration potentielle des estimations issues de l'analyse CTN en présence de nombreux évènements ex-æquos et que ce problème ne soit pas suffisamment pris en compte dans la pratique. En particulier, les chercheurs devraient être conscients des limites possibles des options par défaut utilisées dans les logiciels statistiques disponibles pour mettre en œuvre la régression logistique conditionnelle. Dans R, alors que la fonction *clogit* utilise par défaut la méthode exacte, elle appelle la fonction *coxph* qui utilise par défaut l'approximation d'Efron. Dans SAS, la régression logistique conditionnelle peut être effectuée à l'aide de la procédure PHREG avec une instruction STRATA, qui utilise l'approximation de Breslow par défaut, ou de la procédure LOGISTIC avec une instruction STRATA. Cette dernière procédure s'appuie sur la méthode exacte pour les données discrètes en supposant que les évènements se sont réellement produits exactement au même moment (Allison 2010). Cette procédure est appelée, par exemple, par la macro %nCCsampling qui effectue un échantillonnage de densité d'incidence pour l'analyse CTN (Desai *et al* 2016). Dans nos données simulées, les ex-æquos résultaient d'une mesure imprécise du temps continu. Par conséquent, il convient d'utiliser la méthode exacte en supposant qu'il existe un ordre réel mais inconnu pour les temps d'évènements exæquos. Le calcul de la vraisemblance exacte, qui prend en compte tous les ordres possibles des évènements ex-æquos, peut être chronophage

(Allison 2010) pour les analyses de la cohorte entière mais, selon notre expérience, le temps de calcul était raisonnable pour les analyses CTN.

- **Limites de l'étude**

Dans nos simulations, l'exposition variable dans le temps change au plus une seule fois au cours du suivi et nous avons utilisé le modèle d'exposition actuelle pour la quantifier. Dans la vie réelle, l'exposition médicamenteuse est susceptible de varier plusieurs fois au cours du temps et plusieurs autres modèles d'exposition permettent de mieux prendre en compte cette variation et ainsi de mieux évaluer son association avec un évènement d'intérêt (Abrahamowicz *et al* 2012 ; Gasparrani 2014 ; Pazzagli *et al* 2018 ; Sylvestre & Abrahamowicz 2009).

Les scénarios de simulation considérés nous ont permis d'évaluer uniquement un effet délétère de l'exposition médicamenteuse. Il serait intéressant d'évaluer également les propriétés statistiques de ces schémas d'étude pour l'estimation d'un effet bénéfique qui a son importance en pharmaco-épidémiologie.

Les performances des différentes approches n'ont pas été évaluées avec l'appariement ou l'ajustement sur un ensemble de divers facteurs de confusion fixes et/ou variables dans le temps. L'ajustement ou l'appariement sur plusieurs variables est généralement utilisé pour réduire les biais de confusion. Cependant, les biais ont été observés dans l'étude d'Austin *et al* (2012). malgré l'ajustement des modèles d'analyse CTN sur des covariables fixes dans le temps.

Enfin, nous avons ignoré les risques concurrents dus à d'autres évènements empêchant la survenue de l'évènement d'intérêt.

- **Perspectives du travail**

Bien que notre étude se soit limitée à une exposition variable dans le temps avec un seul changement de "non exposé" à "exposé" pendant le suivi, des travaux ultérieurs devraient étudier des expositions variables dans le temps avec des changements multiples (Desquilbet & Meyer 2005), représentées par des modèles de mesure plus sophistiqués et plus flexibles développés pour mieux prendre en compte la complexité de la relation entre l'exposition médicamenteuse et les évènements de santé. En particulier, le modèle d'exposition cumulée pondérée (ou WCE pour *weighted cumulative exposure,* Sylvestre & Abrahamowicz 2009) présenté dans des études comme étant plus adéquat pour représenter une exposition variable dans le temps (Abrahamowicz *et al* 2012 ; Pazzagli *et al* 2018) peut être utilisé. Les Distributed Lag Non-linear Models (DLNM, Gasparrini 2014), qui généralisent le modèle WCE en combinant deux fonctions pour modéliser de manière flexible la relation exposition-réponse pour l'intensité de l'exposition et la relation temps-réponse pour le temps écoulé depuis l'exposition, peuvent également être utilisés. Une étude évaluant les effets des expositions variables dans le temps avec des rapports de risque décroissants[7] sur la survenue d'un évènement de santé serait intéressant à explorer.

La prise en compte de l'ajustement sur des facteurs de confusion fixes ou variables dans le temps devrait être abordée. Cependant, dans une étude avec des expositions qui varient au cours du temps, les modèles statistiques standard ajustés sur des facteurs de confusion variables dans le temps peuvent conduire à des estimations biaisées de l'effet lorsque ces facteurs de confusion sont affectés par les niveaux d'exposition antérieurs, c'est-à-dire qu'ils agissent comme des intermédiaires dans la relation entre l'exposition et l'évènement d'intérêt (Robins *et al* 2000). Dans ce cas, les modèles structuraux marginaux devraient être utilisés pour ajuster sur les facteurs de confusion variant dans le temps (Pazzagli *et al* 2018 ; Robins *et al* 2000).

Des travaux ultérieurs seraient également nécessaires pour examiner s'il pourrait y avoir un avantage de l'analyse CTN appariée sur des covariables par rapport à l'analyse de cohorte ajustée sur les mêmes covariables tout en tenant compte des facteurs de confusion potentiels. D'autres travaux seraient également justifiés pour comparer les schémas d'étude de cohorte et CTN en présence d'évènements concurrents (Austin *et al* 2012).

4.2 Axe 2

4.2.1 Constitution d'un ensemble de référence

Ce sont 21 méthodes différentes de détection de signaux qui ont été utilisées pour les 57 407 grossesses d'une durée supérieure à 22 semaines d'aménorrhée identifiées dans l'EGB entre 2009 et 2016. Sur un total initial de plus de 800 médicaments de classe ATC (Anatomique, Thérapeutique et Chimique) 5ème niveau, nous en avons retenu exactement 400 qui ont été prescrits au cours d'au moins 5 grossesses. Les 21 méthodes de détection différentes ont abouti à la génération de 81 signaux (au moins par une méthode). Ces 81 signaux ont été classés à partir de l'évaluation clinique de la littérature scientifique publiée et de l'avis d'une experte pharmacovigilante entre les 3 modalités : « Aberrant », « Potentiel » et « Incertain ». Parmi ces 81 signaux, 43 (53.0%) ont été jugés comme « Aberrants » sur un faisceau d'arguments suffisants pour conclure à l'absence de lien avec le risque de prématurité. Pour 24 médicaments, soit 29.7%, les données publiées ne permettaient pas de conclure à la possible présence d'un lien et ont été classés en tant que « Incertain ». Enfin pour 14 (17.3%) d'entre eux, classés comme « Potentiels », des travaux publiés évoquant une possible association avec un risque augmenté de naissance prématurée ont été retrouvés.

La séquence de construction de l'ensemble des associations d'intérêt a donc été inversée : cela a nécessité de définir et de mettre en œuvre les différentes méthodes de détection de signaux, pour restreindre de manière automatisée la liste des associations à catégoriser. Cette réduction a été efficace, permettant de passer de 400 à 81 médicaments, tout en permettant d'obtenir *in fine* des contrôles de nature différente. Elle a été mise en œuvre sur un échantillon au 1/97ème du SNDS ce qui n'invalide pas la conduite future des analyses statistiques sur l'ensemble du SNDS. La catégorisation, légèrement modifiée, a ensuite été menée par une pharmacologue et un clinicien.

Nous avons pu également développer un nouveau critère de détection de signal, le p-RD (p-value relative decrease, décrit dans la section suivante). Sur l'ensemble de référence, ce nouveau critère a montré des performances intéressantes, notamment en termes de réduction des faux positifs, ce qui permet de l'envisager comme outil de détection de signal pour d'autres événements indésirables.

Bien qu'efficace, cette stratégie de construction de l'ensemble de référence, qui repose sur l'expertise malgré tout, reste lourde si on veut l'élargir à d'autres pathologies au-delà de la prématurité. Nous avons donc développé un requêtage automatisé pour 20 pathologies (chez la femme enceinte).

Afin de quantifier de manière automatique l'état des connaissances scientifiques d'une association donnée, nous avons adapté une méthode basée sur les mots-clés des articles scientifiques de MEDLINE. Nous avons utilisé le package RISmed qui permet d'extraire les éléments dans un format directement exploitable sous R. Le langage utilisé dans la classification MeSH n'est pas le même que celui utilisé dans la classification ATC. L'Unified Medical Language System (UMLS) contient un métathésaurus permettant de faire le lien entre différents thésaurus comme le MeSH et la classification ATC. Pour quantifier l'état des connaissances contenues dans les articles, nous avons modifié une méthode initialement développée par Avillach *et al* (JAMIA, 2013) utilisant uniquement le nombre d'articles contenant la combinaison médicament – effet indésirable. Comme nous étudions des pathologies avec des volumes de littérature très disparates, nous avons développé un critère prenant en compte à la fois le nombre total d'articles trouvés pour chaque pathologie et le nombre total de descripteurs (correspondant à des médicaments ou des classes médicamenteuses) mentionnés dans ces articles. Au final, la combinaison du critère automatisé p-RD avec la quantification automatisée de la littérature scientifique pour une association donnée permet de générer des ensembles d'associations d'intérêt qui sont des signaux annotés.

4.2.2 Utilisation des scores pronostiques en grande dimension

Dans des conditions de calcul semblables (mêmes caractéristiques de processeur et paramétrage identique), nous avons montré que le boosting demande plus de temps machine que le lasso. Cette différence est d'autant plus marquée que le nombre de variables augmente. On note une très forte réduction du temps de calcul du boosting quand une présélection est faite avec une division par 5 du temps de calcul. L'effet est également présent pour le lasso mais à une échelle moindre avec une division par 2 du temps de calcul. Nous avons montré que le choix de la méthode d'apprentissage automatique n'avait que très peu d'influence sur le pouvoir prédictif des scores avec des AUC proches pour les deux types de score. Nous avons également montré que la présélection d'un sous ensemble de variables avait un impact négatif sur les AUC du PS mais ne changeait que très peu les AUC du DRS. Alors que les pouvoirs prédictifs des scores sont en médiane peu différents, l'hétérogénéité du PS montre sa forte sensibilité au niveau d'exposition. En comparaison, le DRS se construit avec comme évènement la prématurité chez les non exposées qui reste relativement fréquente. Nous avons comparé différentes utilisations des scores dans le modèle final. L'appariement et la pondération n'ont pas donné

de résultats satisfaisants dans notre étude. Nous avons remarqué que l'algorithme d'appariement avait beaucoup de difficultés à trouver des témoins correspondant suffisamment aux cas en se basant sur le score avec des distances élevées voire des échecs d'appariement amenant à des modélisations impossibles pour certaines expositions rares.

Les associations générées par les différentes méthodes, après correction des p-values pour tests multiples (procédure de Benjamini et Hochberg de contrôle du FDR) ont été évaluées et catégorisées (cf plus haut). C'est le modèle brut qui génère le plus d'associations (76), avec une majorité de faux positifs (57%) et la combinaison hdDRS – boosting – appariement qui génère une seule association (aberrante). La modalité qui génère le plus d'associations après le modèle brut est le DRS construit par boosting et intégré par ajustement. Néanmoins, on observe que cette technique ne suffit pas à réduire suffisamment les faux positifs (23/45 sans pré-sélection). Ce résultat doit néanmoins être nuancé. En effet, l'ensemble des associations utilisées pour l'évaluation est issu d'une sélection par fouille de données, et non sur la base d'hypothèses non exploratoires. Dans ce même travail, nous avons proposé un nouveau critère, le p-RD (p-value relative decrease) basé sur la variation de la p-value observée avec le modèle brut et le modèle intégrant le score. Nous sommes partis de l'idée que s'il existe un effet du médicament sur le risque de prématurité alors la prise en compte du score qui est supposée éliminer une partie de la confusion et des biais devrait permettre une diminution de la p-value entre le modèle brut et le modèle avec le score. A l'inverse, en l'absence d'effet du médicament, on s'attend plutôt à une augmentation de cette dernière. Nous avons appliqué un seuil empirique de 0.20. Avec la modalité précédente, le nombre d'associations retenues était de 12, dont 0 aberrantes. L'utilisation de ce nouveau critère a permis une bien meilleure discrimination en réduisant l'impact du biais d'indication.

Les résultats ont été présentés au congrès de la Société Internationale de Pharmacoépidémiologie (2019, 2020) et publiés dans la revue *Drug Safety* (pour la prématurité).

Ces travaux (2ème axe) ont été menés par Romain Demailly lors de sa thèse dont le cadre était celui de la détection de signaux. Il a pu évaluer de manière très complète les méthodes de prise en compte de la confusion par scores synthétiques en grande dimension, utilisées en pharmacoépidémologie, sur un large ensemble de médicaments (400), dans des conditions réelles, avec pour événement d'intérêt la prématurité. Le travail a été effectué sur l'EGB afin de permettre une évaluation d'un grand nombre de méthodes dans des temps de calcul raisonnables et pour se ménager la possibilité de répliquer des résultats obtenus sur l'EGB dans le SNDS. La constitution d'un ensemble de référence faisait également partie de ses objectifs de thèse.

4.2.3 Analyse par variable instrumentale

Notre cas d'école est l'utilisation d'inhibiteurs sélectifs de la recapture de la sérotonine (ISRS) pendant la grossesse et le risque d'accouchement prématuré (Eke et al. 2016).

- **Population d'analyse et construction de la variable instrumentale**

L'échantillon d'étude est constitué de 22 179 grossesses ayant eu un remboursement d'ISRS au cours de la grossesse par 18 779 prescripteurs différents. Près des trois quarts de ces prescripteurs (13976) l'avait prescrit une seule fois dans cet échantillon (au cours de la grossesses), avec la distribution de la taille de la patientèle suivante (tableau 15) :

Nombre de prescripteurs	Taille de la patientèle
13976	1
3221	2
939	3
332	4
117	5
166	[6 - 10]
24	[11 –20]
4	[21 – 48]

Tableau 15 : distribution de la taille de la patientèle des médecins prescripteurs d'ISRS au cours de la grossesse

Afin de pouvoir construire la VI « Préférence du prescripteur » sans trop réduire la taille de l'échantillon d'analyse, nous avons sélectionné les grossesses dont les prescripteurs correspondants avaient prescrit un ISRS au moins 2 fois (4803 prescripteurs). A l'issue de ce processus, l'échantillon d'analyse contenait 4873 grossesses. Il s'agit d'un premier élargissement de la VI, puisque c'est la préférence du prescripteur au cours de la grossesse, et pas seulement en TBD qui va être construite. D'autres questions se sont posées, avec les choix adoptés suivants pour y répondre :

- si un médecin a prescrit plusieurs types d'ISRS à une même patiente, on considère la première prescription
- si une patiente a eu plusieurs prescripteurs, on garde l'information pour chaque prescripteur : une femme peut contribuer au calcul de la préférence de plusieurs praticiens
- si il y a plusieurs VI pour une femme, on choisit celle du praticien qui a le plus grand nombre de patientes (mais d'autres options ont été explorées).

- **Force de l'instrument**

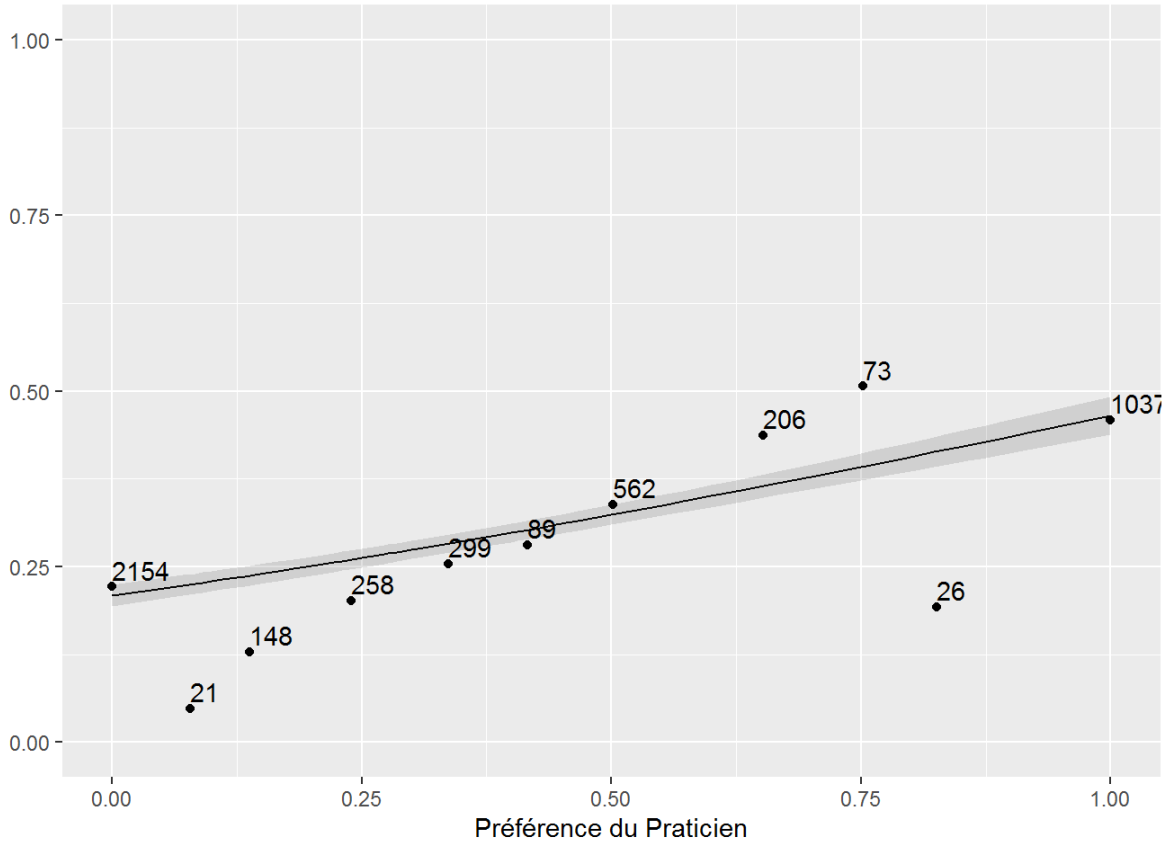
La distribution de la variable instrumentale obtenue est la suivante (tableau 16) :

	Non exposée	Exposée

	N = 3425	N = 1448
Moyenne	0.30	0.49
0	49%	33%
]0 – 0.1]	1%	0
]0.1 – 0.2]	5%	2
]0.2 – 0.3]	4%	3
]0.3 – 0.4]	7%	6
]0.4 – 0.5]	12%	13%
]0.5 – 0.6]	1%	2%
]0.6 – 0.7]	3%	5%
]0.7 – 0.8]	1%	3%
]0.8 – 0.9]	0%	0%
]0.9 – 1]	16%	33%

Tableau 16 : distribution de la VI « Préférence du praticien »

La figure 5 montre qu'il existe une corrélation (0.22) entre la variable instrumentale et l'exposition à T_{BD} :



.Figure 5 : Probabilité de traitement (escitalopram) en fonction de la préférence du médecin, avec indication des effectifs (toutes les grossesses ont été traitées avec un ISRS).

- **Résultats des différents modèles**

Nous avons appliqué notre stratégie d'analyse allant des modèles bruts à des modèles ajustés sur plusieurs covariables (tableau 17), pour les analyses classiques et pour les analyses par variable instrumentale. En ce qui concerne les modèles classiques, parmi ces covariables, l'âge de la mère est un facteur de risque pour lequel nous avons pu observer sur nos données la courbe en U du risque de prématurité induit par l'âge de la mère (figure 6), sans toutefois retrouver cette association significative ; nous retrouvons une association significative avec la CMU-C, mais pas avec FDep (résultats non détaillés).

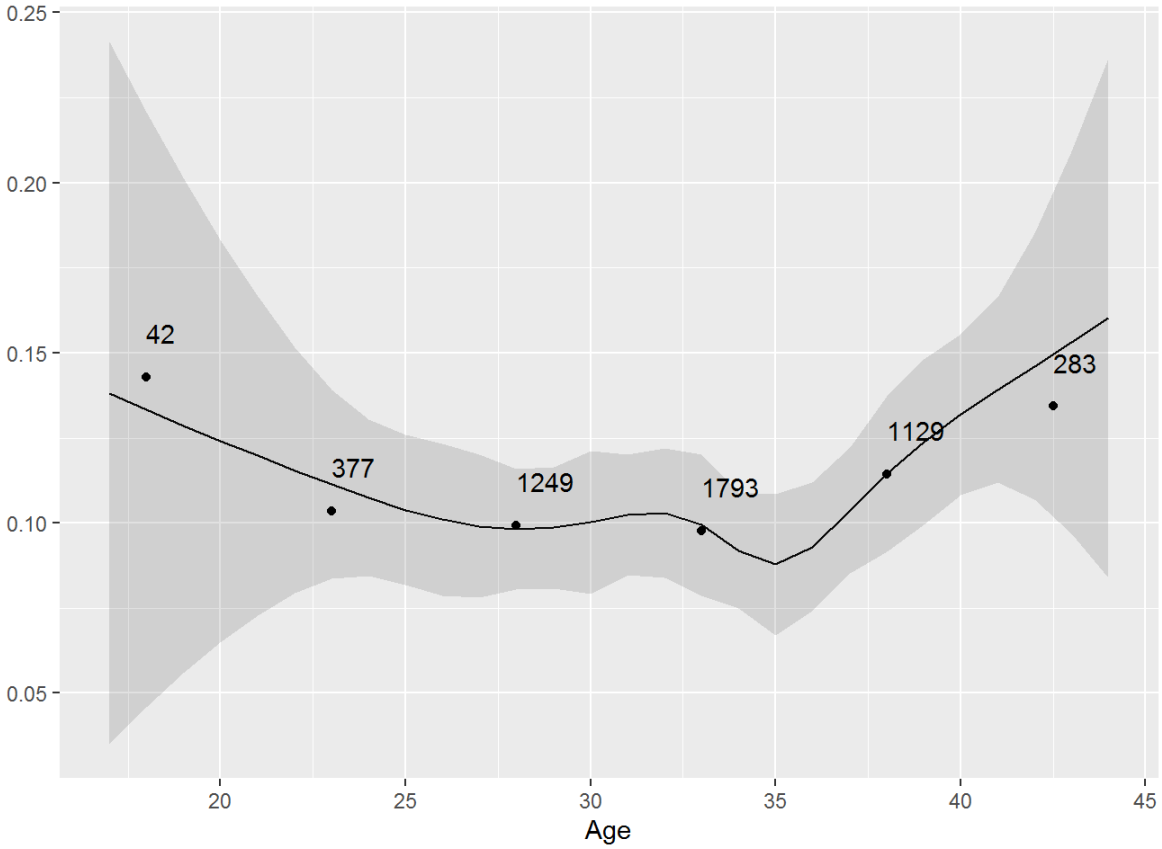


Figure 6 : risque de prématurité en fonction de l'âge de la mère, avec indication des effectifs (toutes les grossesses ont été traitées avec un ISRS)

L'analyse classique, brute ou ajustée, conclut à l'existence d'un risque de prématurité induit par l'escitalopram versus les autres antidépresseurs de la classe des ISRS. Ce risque est augmenté lorsqu'on introduit les covariables. Un calcul de E-value (VanderWeele 2016), donne, pour l'estimation brute une valeur de 1.8 (1.1 pour l'intervalle de confiance), et pour l'analyse ajustée, 2.1 (1.1 pour l'intervalle de confiance) : c'est la force minimale de l'association que devrait avoir un facteur de confusion non mesuré avec l'exposition et la prématurité, conditionnellement aux covariables pour complètement expliquer l'association trouvée. Un tel facteur de confusion n'est pas à exclure, ce qui justifie de tenter des approches de correction des biais.

L'approche par variable instrumentale n'apporte pas de réelle modification de l'estimation de l'association (plutôt augmentée), mais on observe un énorme élargissement de l'intervalle de confiance, qui ne conduit plus à une association significative. L'analyse par VI ne permet pas de conclure.

Méthode	**OR**	**IC**
Brute	**1.23**	**[1.01 – 1.49]**

Classique ajustée : âge, exposition T1, cmuc, Fdep	**1.38**	**[1.01 – 1.86]**
VI sans covariable	**1.15**	**[0.44 – 2.75]**
VI, âge, exposition T1, cmuc, Fdep	**1.54**	**[0.17 – 14.69]**
VI, âge, exposion T1, cmuc, Fdep, Spécialité	**1.42**	**[0.16 – 13.03]**

Tableau 17 : analyses classiques et par variable instrumentale de l'association entre prématurité et exposition à l'escitalopram au cours du trimestre précédent l'accouchement versus les autres antidépresseurs de la classe des ISRS

- **Discussion et perspectives**

Le risque de prématurité étudié avec les données du SNDS est un cas d'école propice à identifier les difficultés liées à la confusion non mesurée. On sait que plusieurs facteurs de risque de prématurité ne sont pas mesurés (infection, antécédent de prématurité par exemple). On ne connaît pas vraiment ce que pourraient être les facteurs de confusion. Avant ce projet, il n'y avait pas de « gold standard » permettant de confronter les résultats obtenus avec différentes approches à ces gold standards. La construction d'un ensemble de référence pour la prématurité est donc une contribution importante : elle a déjà permis des évaluations méthodologiques, dans le cadre de la présente recherche, mais aussi pour une autre thèse encadrée dans l'équipe qui a démarré et qui considère d'autres approches pour les facteurs de confusion non mesurés. Elle s'est appuyée sur l'EGB, ce qui fait que les médicaments peu fréquents ne peuvent pas être considérés. Par ailleurs, le SNDS (et donc l'EGB) ne permet d'étudier que les médicaments remboursés. Il s'agit d'une étape encore imparfaite (on parle de signaux intéressants, ou écartés), mais qui améliore la conduite des recherches, et enrichit les informations accessibles aux cliniciens, mais aussi l'aide à la décision de santé publique.

L'évaluation à grande échelle de l'utilisation (sous toutes ses formes) des scores synthétiques en grande dimension a montré qu'ils ne pouvaient à eux seuls suffire à réduire drastiquement les biais de confusion. De nombreux paramètres sont à définir pour leur calcul. Il est également envisageable que de bonnes performances pourraient être obtenues en ne se limitant pas à une seule approche mais en combinant de nombreux "apprenants", comme le proposent Ju *et al.* (Ju *et al*. 2017) avec le Super Learner (van der Laan *et al.* 2007). Évidemment, la charge de calcul est un paramètre important à prendre en compte lors de la mise en œuvre de tels outils sur de très grands ensembles de données. Par ailleurs, la proposition d'un indicateur dérivé (p-RD), dans le cadre de la détection de signal pourrait être prometteuse en pharmacoépidémiologie. Un autre point qui n'a pas été abordé avec le hdPS concerne sa dérivation lorsque l'exposition d'intérêt varie dans le temps ; une perspective porte sur l'adaptation de la méthodologie du score de propension pour tenir compte de l'exposition dépendant du temps.

En ce qui concerne l'utilisation des variables instrumentales, celle-ci pose de nombreuses difficultés : tout d'abord, la nécessité d'un comparateur (un autre médicament), la construction même de la préférence du médecin diminuent la taille de l'échantillon analysé, ce qui pose un problème de puissance de l'analyse, car par ailleurs, les méthodes d'analyse par VI requièrent de larges effectifs. Une perspective est d'élargir l'analyse en incluant d'autres années. Une autre possibilité est de modifier la construction de la préférence du médecin : par exemple en élargissant son calcul à toutes les femmes en âge d'être enceinte. L'instrument construit est peut-être trop faible, et par ailleurs, on ne peut pas vérifier les hypothèses nécessaires à l'analyse, qui si elles sont non respectées pourraient amplifier les biais.

5. Difficultés rencontrées

Deux doctorants ont travaillé sur chacun des deux axes du projet. La crise sanitaire a énormément impacté la conduite des recherches, mais aussi des événements de grossesse (des deux côtés), ce qui fait que certains aspects n'ont pas pu être abordés, et toute la valorisation qui était prévue n'a pas complètement abouti. En ce qui concerne le 2ème axe, sur la partie des variables instrumentales nous avons élargi le projet à une ingénieur de recherche biostatisticienne qui fait partie du pôle méthodologique du CESP, ce qui a eu comme effet favorable, en plus de pouvoir avancer sur le sujet, d'amener une collaboration transversale au sein du centre.

References

Abrahamowicz, M., Beauchamp, M.E. & Sylvestre, M.P., 2012. Comparison of alternative models for linking drug exposure with adverse effects. *Stat Med*, 31(11–12), pp.1014–1030.

Allison PD. *Survival Analysis using SAS: a practical guide*. Cary, NC: Sas Institute Inc., 2010.

Austin PC, Anderson GM, Cigsar C, Gruneir A. Comparing the cohort design and the nested case-control design in the presence of both time-invariant and time-dependent treatment and competing risks: bias and precision. *Pharmacoepidemiol Drug Saf* 2012;21:714-724.

Avillach P, Dufour J-C, Diallo G, Salvo F, Joubert M, Thiessard F, et al. Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: a contribution from the EU-ADR project. J Am Med Inform Assoc. 2013 May 1;20(3):446–52.

Bertke S, Hein M, Schubauer-Berigan M, Deddens J. A simulation study of relative efficiency and bias in the nested case-control study design. *Epidemiol Methods* 2013;2:85-93.

Billioti de Gage S, Begaud B, Bazin F et al. Benzodiazepine use and risk of dementia: prospective population based study. *BMJ* 2012;345:e6231.

Breslow N. Covariance analysis of censored survival data. *Biometrics* 1974;30:89-99.

Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. *J Am Stat Assoc* 1983;78:1-12.

Brickner CP. Estimating the relationship between a transient effect and the onset of an acute event: a comparison of the case-crossover design and cohort design, Rutgers: State University of New Jersey 2015.

Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology.* 2006;17(3):268–75.

Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models. Monographs on Statistics and Applied Probability.* London: Chapman & Hall; 1995.

Cheung YB, Ma X, Lam KF, Li J, Milligan P. Bias control in the analysis of case-control studies with incidence density sampling. *Int J Epidemiol* 2019;48:1981-1991.

Demailly, R. et al., 2017. Prescription drug use during pregnancy in France: a study from the national health insurance permanent sample. *Pharmacoepidemiology and drug safety*, 26(9), pp.1126–1134.

Desai RJ, Glynn RJ, Wang S, Gagne JJ. Performance of Disease Risk Score Matching in Nested Case-Control Studies: A Simulation Study. *Am J Epidemiol* 2016;183:949-57.

Desquilbet L, Meyer L. Variables dépendantes du temps dans le modèle de Cox: théorie et pratique [Time-dependent covariates in the Cox proportional hazards model. Theory and practice]. *Rev Epidemiol Sante Publique* 2005;53:51-68.

Efron B. The efficiency of Cox's likelihood function for censored data. *J Am Stat Assoc* 1977;72:557-565.

Eke, A.C., Saccone, G. & Berghella, V., 2016. Selective serotonin reuptake inhibitor (SSRI) use during pregnancy and risk of preterm birth: a systematic review and meta-analysis. *BJOG : an international journal of obstetrics and gynaecology*, 123(12), pp.1900–1907.

Ertefaie, A. et al., 2017. A tutorial on the use of instrumental variables in pharmacoepidemiology. *Pharmacoepidemiology and Drug Safety*, 26(4), pp.357–367.

Essebag V, Platt RW, Abrahamowicz M, Pilote L. Comparison of nested case-control and survival analysis methodologies for analysis of time-dependent exposure. *BMC Med Res Methodol* 2005;5:5.

Farewell VT, Prentice RL. The approximation of partial likelihood with emphasis on case-control studies. *Biometrika* 1980;67:273-278.

Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. Biometrics 1995 Mar;51(1):228-35

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika* 1993;80:27-38.

Franklin, J.M. et al., 2017. Comparing the performance of propensity score methods in healthcare database studies with rare outcomes. *Statistics in Medicine*, 36(12), pp.1946–1963.

Franklin, J.M. et al., 2015. Regularized Regression Versus the High-Dimensional Propensity Score for Confounding Adjustment in Secondary Database Analyses. *Am J Epidemiol*, 182(7), pp.651–659.

Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med* 2014;33:881-899.

Goldstein L, Langholz B. Asymptotic theory for nested case-control sampling in the Cox regression model. *Ann Stat* 1992;20:1903-1928.

Goueslard K, Cottenet J, Benzenine E, Tubert-Bitter P, Quantin C. Validation study: evaluation of the metrological quality of French hospital data for perinatal algorithms. BMJ Open. 2020 May 12;10(5):e035218. doi: 10.1136/bmjopen-2019-035218.

Greene WH. *Econometric Analysis, 7th edition, international edition.* Boston: Pearson; 2012.

Hak E, Wei F, Grobbee DE, Nichol KL. A nested case-control study of influenza vaccination was a cost-effective alternative to a full cohort analysis. *J Clin Epidemiol* 2004;57:875-80.

Hanf M, Quantin C, Farrington P, Benzenine E, Hocine NM, Velten M, Tubert-Bitter P, Escolano S. Validation of the French national health insurance information system as a tool in vaccine safety assessment: application to febrile convulsions after pediatric measles/mumps/rubella immunization Vaccine. 2013 Dec 2;31(49):5856-62.

Hansen BB. The prognostic analogue of the propensity score. Biometrika. 2008 Feb 4;95(2):481–8.

Hernán MA, Jick SS, Olek MJ, Jick H. Recombinant hepatitis B vaccine and the risk of multiple sclerosis: a prospective study. *Neurology* 2004;63:838-42.

Koladjo BF, Escolano S, Tubert-Bitter P. Instrumental variable analysis in the context of dichotomous outcome and exposure with a numerical experiment in pharmacoepidemiology. BMC Med Res Methodol. 2018 Jun 22;18(1):61.

Ju, C. et al., 2017. Propensity score prediction for electronic healthcare databases using Super Learner and High-dimensional Propensity Score Methods.

van der Laan, M.J., Polley, E.C. & Hubbard, A.E., 2007. Super Learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1).

Langholz B, Richardson DB. Fitting general relative risk models for survival time and matched case-control analysis. *Am J Epidemiol* 2010;171:377-83.

Lee, B.K., Lessler, J. & Stuart, E.A., 2010. Improving propensity score weighting using machine

learning. *Statistics in medicine*, 29(3), pp.337–46.

Leffondré K, Abrahamowicz M, Siemiatycki J. Evaluation of Cox's model and logistic regression for matched case-control data with time-dependent covariates: a simulation study. *Stat Med* 2003;22:3781-94.

Liddell FDK, McDonald JC, Thomas DC, Cunliffe SV. Methods of cohort analysis: appraisal by application to asbestos mining. *J Royal Stat Soc* 1977;A140:469.

Nestler S. Using instrumental variables to estimate the parameters in unconditional and conditional second-order latent growth models. *Struct Equ Model A Multidiscip J.* 2015;22(3):461–73.

Pang D. A relative power table for nested matched case-control studies. *Occup Environ Med* 1999;56:67-9.

Pazzagli L, Linder M, Zhang M et al. Methods for time-varying exposure related problems in pharmacoepidemiology: An overview. *Pharmacoepidemiol Drug Saf* 2018;27:148-160.

Peto R. Contribution to discussion of paper: Regression models and lifetables, by D.R. Cox. *J Stat Royal Stat Soc* 1972;B34:205-207.

Quantin C, Yamdjieu Ngadeu C, Cottenet J, Escolano S, Bechraoui-Quantin S, Rozenberg P, Tubert-Bitter P, Gouyon JB. Early exposure of pregnant women to non-steroidal anti-inflammatory drugs delivered outside hospitals and preterm birth risk: nationwide cohort study BJOG. 2021 Sep;128(10):1575-1584.

Rassen, J.A. et al., 2009. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *American journal of epidemiology*, 169(3), pp.273–84.

Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550-560.

Schneeweiss, S. et al., 2009. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology*, 20(4), pp.512–522.

Schneeweiss, S. et al., 2017. Variable Selection for Confounding Adjustment in High-dimensional Covariate Spaces When Analyzing Healthcare Databases. *Epidemiology*, 28(2), pp.237–248.

Sun JX, Sinha S, Wang S, Maiti T. Bias reduction in conditional logistic regression. *Stat Med* 2011;30:348-55.

Sylvestre MP, Abrahamowicz M. Comparison of algorithms to generate event times conditional on time-dependent covariates. *Stat Med* 2008;27:2618-2634.

Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat Med* 2009;28:3437-3453.

Thiébaut ACM, Benichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. *Stat Med* 2004;23:3803-3820.

Weill, A., Rudant, J. & Coste, J., 2017. Utilisation des données de l'Assurance maladie française pour

étudier la qualité des pratiques et la sécurité des soins dans le domaine du médicament : revue de 204 articles publiés entre 2007 et 2016. *Revue d'Épidémiologie et de Santé Publique*, 65, p.S22.

VanderWeele TJ & Ding P (2017). Sensitivity analysis in observational research: introducing the E-value. Annals of Internal Medicine, 167(4), 268-274.

DIFFUSION/COMMUNICATION

Ces livrables peuvent faire l'objet de communications externes

Articles

Manitchoko L, Abrahamowicz M, Tubert-Bitter P, Benichou J*, Thiébaut ACM Comparison of cohort and nested case-control designs for estimating the effect of time-varying drug exposure on the risk of adverse event in the presence of ties. Biom J. 2023 Feb 27:e2100384. doi: 10.1002/bimj.202100384. Online ahead of print. PMID: 36846937. (*Contribution égale) Accès libre

https://onlinelibrary.wiley.com/doi/10.1002/bimj.202100384

Demailly R, Escolano S, Haramburu F, Tubert-Bitter P, Ahmed I. Identifying Drugs Inducing Prematurity by Mining Claims Data with High-Dimensional Confounder Score Strategies**.** Drug Saf. 2020 Jun;43(6):549-559. doi: 10.1007/s40264-020-00916-5.PMID: 32124266. Joint

Thèses

Liliane Manitchoko a soutenu sa thèse en biostatistique « Modélisation des expositions médicamenteuses et de leurs effets variant dans le temps : comparaison de méthodes d'analyse statistique » le 1er juillet 2022 (doctorat de l'Université Paris Saclay). https://theses.fr/2022UPASR008

Cette thèse a été financée à l'aide de subventions de l'ANSM puis de l'IReSP et de vacations sur budget propre de l'équipe. Liliane Manitchoko est actuellement en formation.

Romain Demailly a soutenu sa thèse en épidémiologie « Détection automatisée de signaux en pharmacovigilance chez la femme enceinte à partir de bases médico-administratives » le 3 mars 2020 (doctorat de l'Université Paris Saclay). Il a été nommé maître de conférences à la faculté de médecine et maïeutique de l'université catholique de Lille et poursuit sa collaboration avec nous comme chercheur associé.

Communications

- Communications orales dans des conférences nationales ou réunions scientifiques

Antoni G. Prise en compte de la confusion cachée par l'utilisation de variable instrumentale ; application à l'estimation du risque de prématurité chez la femme enceinte exposée à l'escitalopram à partir des données du SNDS. Séminaire interne CESP, Villejuif (France). 31 mars 2023

Manitchoko L, Bénichou J, Thiébaut A. Comparison of cohort and nested case-control designs for estimating time-varying drug exposure effect on adverse event risk. Séminaire de recherche hebdomadaire du Bureau de Biostatistique et d'Epidémiologie de Gustave Roussy et de l'équipe ONCOSTAT du Centre de recherche en Epidémiologie et Santé des Population (UMR 1018, Inserm, Université Paris-Saclay, UVSQ), Villejuif, 29 novembre 2021.

Manitchoko L, Bénichou J, Thiébaut A. Comparaison de méthodes d'estimation de l'effet d'une exposition médicamenteuse sur le risque d'évènement indésirable. Journées du GDR « Statistiques et santé », 10 – 11 octobre 2019, Paris (oral)

- Communications dans des conférences internationales

Manitchoko L, Abrahamowicz M, Tubert-Bitter P, Bénichou J, Thiébaut A. Comparison of cohort and nested case-control designs for estimating the effect of time-varying drug exposures. 31st International Biometric Conference, Riga, 10-15 juillet 2022. (oral)

Manitchoko L, Bénichou J, Thiébaut A. Comparison of statistical methods for estimating time-varying treatment effect on adverse event risk. 42nd annual conference of the International Society for Clinical Biostatistics, Lyon, 18-22 juillet 2021. (oral)

Manitchoko L, Bénichou J, Thiébaut A. Comparison of cohort and nested case- control analysis for estimating the effect of time-varying treatment on the risk of adverse event. Channel Network Conference, Paris, 7-9 avril 2021. (poster)

Demailly R, Escolano S, Tubert-Bitter P, Ahmed I. Automated Large-scale Signal Detection In Pregnant Women From Medico-administrative Databases And Scientific Literature**.** 36th International Conference on Pharmacoepidemiology (ICPE), à l'origine Aug 26-30 2020, Berlin, Germany (poster). Online event Sept 16-17 2020

Demailly R, Escolano S, Haramburu F, Tubert-Bitter P, Ahmed I. Comparing The Use Of Propensity And Disease Risk Scores In Signal Detection For Prematurity On Medico-administrative Database. 35th ICPE, Aug 24-28 2019, Philadelphia, USA (oral)

Liste des séminaires ou colloques en rapport avec le projet financé auxquels vous avez participé et/ou organisé durant la période (et des missions à l'étranger)

(*Précisez la date, le lieu, l'objet, le nombre des participants)*

35th ICPE, Aug 24-28 2019, Philadelphia, USA

Journées du GDR « Statistiques et santé », 10 – 11 octobre 2019, Paris

36th International Conference on Pharmacoepidemiology (ICPE), à l'origine Aug 26-30 2020, Berlin, Germany (poster). Online event Sept 16-17 2020

Channel Network Conference, Paris, 7-9 avril 2021

42nd annual conference of the International Society for Clinical Biostatistics, Lyon, 18-22 juillet 2021

31st International Biometric Conference, Riga, 10-15 juillet 2022

Liste des éventuelles missions à l'étranger effectuées dans le cadre du projet

Liste des communications au grand public

Publications, site internet, interview, questionnaire, forum, plaquette de présentation

Glossaire

Livrable : tout composant matérialisant le résultat de la prestation de réalisation. Toute production émise par le titulaire au cours du projet : document, courrier revêtant un caractère officiel, module de code logiciel, dossiers de tests, application intégrée, objet, dispositif…

Livrable interne : réalisé au sein du programme et non communiqué à l'extérieur du programme.

Livrable externe : élément diffusé ou livré hors de la communauté du projet de recherche.

Faits marquants : élément non nécessairement quantifiable mais significatif pour le projet.

Signature du rédacteur et/ou du coordinateur scientifique de projet (si différents) :

Ce document est à renvoyer aux adresses suivantes :

suiviprojets.iresp@inserm.fr