Article original / Original Article

Journal de Gestion et d'Économie Médicales
2014, Vol. 32, n° 5-6, 371-388

Évaluation quasi-expérimentale des interventions médicales : méthode des variables instrumentales

Salah Ghabri[1], Robert Launois[2]

[1] Service d'Évaluation Économique et de Santé Publique (SEESP), Haute Autorité de Santé

[2] Réseau d'Évaluation en Économie de la Santé (REES), 28 Rue d'Assas, 75006 Paris

Correspondance :
Salah Ghabri, Haute Autorité de Santé, 2 Avenue du Stade de France,
93218 Saint-Denis La Plaine Saint Denis
E_mail : s.ghabri@has-sante.fr

RÉSUMÉ

L'objectif de cet article est de présenter, sans formalisation statistique excessive, la méthode des variables instrumentales (VI) pour analyser l'effet causal de l'efficacité et de la sécurité comparative des traitements dans des conditions en vie réelle. La méthode des VI est utilisée dans le cadre des études observationnelles. Nous l'illustrerons à l'aide d'un exemple de la littérature portant sur la sécurité comparative des anti-inflammatoires non stéroïdiens (AINS) utilisant les préférences de prescription des médecins pour les AINS comme VI. L'approche des VI permet de contrôler le biais d'indication mais elle est toutefois très contraignante au niveau de la vérification de la validité des VI.

Mots-clefs : Variable instrumentale, confusion par indication, efficacité comparative, étude observationnelle, préférences de prescription des médecins.

ABSTRACT

QUASI-EXPERIMENTAL EVALUATION OF MEDICAL INTERVENTIONS: INSTRUMENTAL VARIABLES METHOD

The aim of this paper is to present, without excessive statistical formalisation, the method of instrumental variables (IV) for making causal inference in comparative effectiveness and safety studies of medical interventions. The IV method is used in observational studies. We illustrate our paper with an example from the literature on the comparative safety of non-steroidal anti-inflammatory drugs (NSAID) using physician prescribing preferences for NSAID as an instrumental variable. The IV method allows controlling for confounding by indication. However, testing the validity of IV is challenging.

***Keywords**: Instrumental variables, confounding by indication, comparative effectiveness research, observational study, physician prescribing preferences.*

Conflits d'intérêts : aucun

INTRODUCTION

L'analyse comparative en vie réelle de l'efficacité ou de la sécurité des interventions médicales suscite un intérêt croissant de la part des autorités de santé. Deux causes expliquent le développement de la littérature abondante qui a été consacrée au sujet depuis le début des années 2000 : le développement de nouvelles méthodes pour mesurer quel pouvait être l'effet propre d'un nouveau traitement dans le cadre d'études non expérimentales, et l'élargissement des possibilités d'accès à des banques de données médico-administratives de grande dimension. En France, la problématique est plus que jamais d'actualité avec la signature récente de l'accord cadre LEEM-CEPS qui introduit la notion de prix conditionnels dont la pérennisation sera subordonnée aux résultats observés en situation réelle [1].

Le plan d'expérience idéal pour analyser l'efficacité comparative est l'essai randomisé en double aveugle avec groupe contrôle. Néanmoins, pour de multiples raisons (faisabilité, problèmes éthiques ou financiers), il est parfois difficile de mettre en œuvre un tel schéma d'étude. L'absence de randomisation rend difficile la mise en évidence du lien de cause à effet entre le traitement administré et le résultat obtenu. Les différences observées dans les résultats entre les groupes peuvent être soit liées au traitement soit dues à des différences entre les caractéristiques des populations comparées. Lorsque tel est le cas, l'étude est entachée de biais.

Deux catégories de biais caractérisent les données observationnelles : le biais lié à des variables observables qualifié de biais « mesurable » (*overt bias*) et celui associé à des variables non observables (ou omises) appelé biais caché (*hidden bias)*. Le biais observable se traduit par des facteurs de confusion qui sont à la fois liés à l'allocation (ou choix) du traitement (T) et au critère de jugement (Y). Il peut alors être contrôlé en ayant notamment recours à la méthode du score de propension (Rosenbaum et Rubin, 1983) [2] et/ou la méthode de régression standard (régression linéaire, régression logistique, régression de Cox...). Cependant le problème du contrôle des biais des variables non observables reste entier. Deux techniques peuvent être envisagées pour en assurer la maîtrise : les modèles de sélection de Heckman [3] et l'approche des variables instrumentales [4,5].

L'approche des variables instrumentales est très utilisée par les économistes dans les domaines de la finance et de la consommation des substances licites et/ou illicites [6]. Le travail d'Angrist et al. [4] a initialement introduit la méthode des variables instrumentales dans les essais cliniques pour étudier l'observance. En dépit de quelques travaux en épidémiologie [7,8], son application dans l'évaluation des traitements mis en œuvre dans la vraie vie est assez récente. À partir de la seconde moitié des années 2000, la littérature sur les variables instrumentales s'est progressivement développée. La technique a été employée pour évaluer les effets du médicament en vie réelle [9,10,11] et pour étudier l'impact du cathétérisme cardiaque sur le taux de mortalité des patients souffrant d'infarctus aigu du myocarde aux États-Unis [12].

L'objectif de cet article est de présenter, sans formalisation statistique excessive, comment la méthode des variables instrumentales est mise en œuvre dans le cadre des études observationnelles pour analyser la sécurité comparative des médicaments en vie réelle [13] ; d'enrichir la réflexion sur la possibilité d'exploiter des bases de données administratives françaises après la signature récente de l'accord cadre LEEM-CEPS qui introduit la notion de prix conditionnels dont la pérennisation sera subordonnée aux résultats observés en situation réelle, d'expliquer ses modalités d'utilisation à la lumière d'exemples empruntés à la littérature récente et d'en montrer les limites [14, 15, 16, 17].

La première section rappellera brièvement comment l'effet causal peut être rigoureusement isolé dans un contexte expérimental. La deuxième section discutera les biais induits par les différents schémas d'études observationnelles. La troisième section illustrera sous forme de diagramme l'approche des variables instrumentales en expliquant son principe, et en montrant ses avantages et ses limites. La quatrième section donnera un exemple d'utilisation de cette technique dans le cas de la prescription des anti-inflammatoires non stéroïdiens (AINS) et mettra l'accent sur l'intérêt des travaux de Brookhart et al. [9,10, 16] et de Rassen et al. [18] portant sur des données américaines et le travail de Davies et al [19] portant sur les données britanniques. La conclusion résumera les principaux enseignements et les perspectives de développement de cette méthode.

CONTEXTE EXPÉRIMENTAL ET EFFET CAUSAL

Un essai randomisé en double aveugle avec groupe de contrôle est le plan d'expérience idéal pour analyser l'efficacité d'un traitement. La randomisation rend les groupes comparables au début de l'étude en équilibrant les caractéristiques des sujets affectés à chacun des groupes, que ces caractéristiques soient observables ou non. Un risque de biais d'attrition ne peut cependant pas être complètement exclu lorsqu'il existe des perdus de vue du fait de non réponse ou d'un défaut d'observance. En son absence, le double aveugle

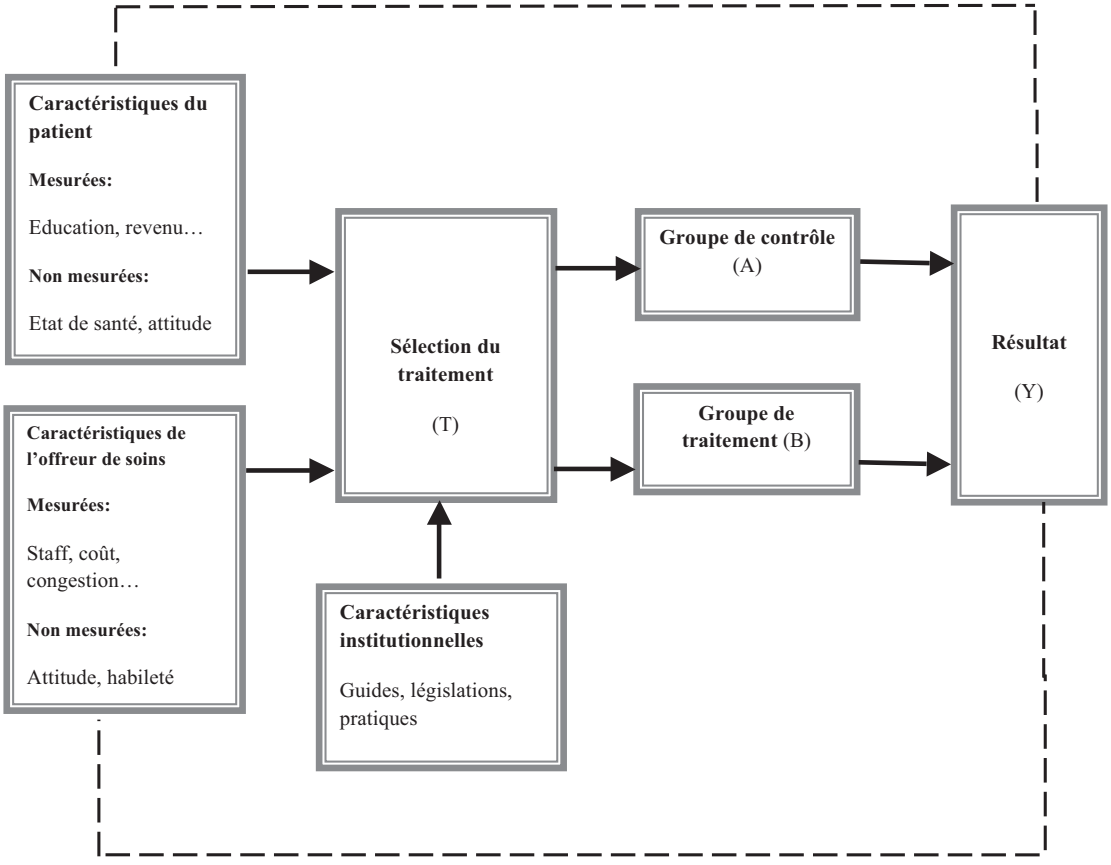
permet de neutraliser l'effet patient ainsi que l'effet investigateur.

Dans ce contexte, les différences observées sur le critère de jugement entre les groupes sont attribuables à l'effet du seul traitement.

ÉTUDES OBSERVATIONNELLES ET BIAIS

Pour des raisons éthiques ou opérationnelles relatives à la mise en place d'un plan d'expérience randomisé, il n'est pas toujours facile de réaliser des essais randomisés lorsqu'on cherche à conduire une étude de médicament ou d'interventions médicales en vie réelle. Des études observationnelles sont souvent mises en place. En l'absence d'affectation aléatoire, une association peut exister non seulement entre les facteurs de confusion (observables/non observables) et le choix du traitement, mais aussi entre les facteurs de confusion et le critère de jugement. Dans les études observationnelles (figure1), les caractéristiques initiales de l'état de santé du patient sont une source potentielle de biais, (exemple, l'intensité de la douleur) puisqu'elles sont à la fois liées au choix du traitement et à son critère de jugement. En pratique, nous observons deux cas fréquents dans les études observationnelles, que nous allons expliciter. Lorsqu'il s'agit d'études prospectives de terrain, les variables

Figure 1 : Schéma standard d'une étude observationnelle – Source : Pizer (2009)



d'intérêt sont colligées. En revanche, dans les banques de données (ou les registres) un grand nombre de variables ne sont pas documentées.

1. Schéma d'étude avec collection de la plupart des variables d'intérêt

Le premier schéma d'une étude observationnelle idéalement préféré par le chercheur est celui où la plupart des variables pertinentes (par exemple, toutes les caractéristiques individuelles importantes au regard du phénomène étudié) ont été identifiées dans le protocole de l'étude et ont été présentées dans les données de l'enquête en question. Elle illustre une situation où les biais sont « mesurables ». Les méthodes classiques : score de propension et régressions standards (linéaires, logistiques, modèles de survie ou de comptage) permettent de les neutraliser. L'approche des scores de propension[1] permet de rendre les populations incluses dans chacun des groupes comparables en équilibrant leurs caractéristiques en début de l'étude, et l'analyse de régression multivariée classique permet (en présence ou en absence des scores de propension) d'ajuster le critère de jugement sur d'éventuelles covariables.

2. Études utilisant des banques de données administratives et variables non observables

Un second schéma d'une étude observationnelle est le cas d'une base de données très volumineuse (registres, banques de données de l'assurance maladie…) où un grand nombre des variables ne sont pas documentées. Les biais observables peuvent être contrôlés à l'aide des méthodes précédemment décrites. Néanmoins, la corrélation entre des variables inobservables avec le traitement et le critère de jugement risque d'induire des biais *« cachés»* que les approches de régressions classiques et/ou les scores de propension sont incapables de contrôler. Ces biais cachés sont liés à l'existence de variables importantes qui ont été omises au départ comme, par exemple, les habitudes de prescription des médecins, les choix thérapeutiques, l'organisation des soins dans un établissement de santé, les plaintes des patients quant à la sévérité de leur douleur…

Une solution pour neutraliser ces biais cachés, consiste à chercher une variable dite « instrument » qui exerce une influence sur le choix du traitement mais non sur son critère de jugement. On s'intéresse aux personnes qui ont bénéficié d'un traitement parce qu'elles ont été exposées à une certaine valeur de la variable instrumentale mais qui n'en n'auraient pas bénéficié si celle-ci avait pris une autre valeur.

LA MÉTHODE DE LA VARIABLE INSTRUMENTALE (VI)

Les variables instrumentales permettent de contrôler les biais des variables non observables ou omises (biais de sélection...). Elles sont très utiles pour évaluer les traitements médicaux en vie réelle. La méthode recherche une source de variation qui joue le rôle d'un tirage au sort dans une étude expérimentale, où les patients à priori suffisamment comparables sont affectés à des traitements différents. Un schéma d'étude doit justifier le choix de la variable instrumentale et décrire le support sur lequel on identifie un effet causal dans le choix du traitement.

[1] Le score de propension est la probabilité pour un patient de caractéristiques données (co-variables observées) d'être exposé à un traitement. Chaque patient exposé est apparié au patient non exposé dont le score est le plus proche. Le sous-échantillon ainsi constitué a des caractéristiques proches de l'essai clinique randomisé (pas de différence entre les groupes des patients exposés et non exposés à l'étape initiale de l'étude).

1. Analyse par la méthode VI

L'approche de la variable instrumentale se déroule en deux étapes :

Étape 1 : La recherche de la variable instrumentale

Un instrument *« valide » est une variable qui* :

Condition 1 : est fortement liée au fait de recevoir du traitement ;

Condition 2 : n'a pas d'effet sur le critère de jugement excepté à travers le traitement *(condition d'exclusion)*[2] ;

Condition 3 : qui n'est pas liée aux caractéristiques du patient.

La figure 2 illustre le concept de variable instrumentale et les trois conditions qui doivent être respectées pour qu'il puisse être mis en œuvre. Si ces conditions sont remplies, l'estimation par la méthode VI de l'effet du traitement sur le critère de jugement Y est asymptotiquement non biaisée [4].

Étape 2 : Calcul et modélisation économétrique de la variable instrumentale

Conformément à l'objectif pédagogique de cet article, nous nous focalisons sur le cas simple d'une seule variable instrumentale qui caractérise non seulement la littérature actuelle sur les études de post-inscription mais qui permet aussi de montrer que le choix de la variable instrumentale doit être préalablement justifié d'un point de vue théorique avant toute vérification empirique confortant son utilisation. Dans le cas d'une seule variable instrumentale dichotomique, l'effet du traitement est donné par la formule de Wald [15] :

$$Effet\ VI = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{Pr[T \mid Z = 1] - Pr[T \mid Z = 0]}$$

Où Y est le critère de jugement, T est le traitement, Z est la variable instrumentale, E[Y|Z=i] est l'effet du traitement conditionnellement à la valeur 1 ou 0 de la variable instrumentale. Pr (T|Z=i) est la probabilité de recevoir le traitement T conditionnellement à la valeur i de la variable instrumentale Z(i=0 ou1).

Si l'instrument est parfait, nous sommes dans le cas d'un schéma expérimental randomisé, l'effet du traitement est équivalent à celui de l'intention de traiter. Dans ce cas (i) : 100% des patients où Z=1 sont traités par T=1 et aucun de ceux avec Z=0. D'où (ii) : Pr {T|Z=1} – Pr {T|Z=0}=1, ce qui conduit à (iii) : B_{VI}=E[Y|Z=1]-E[Y|Z=0]=B_{ITT}

L'annexe 1 décrit la modélisation économétrique des variables instrumentales (variantes linéaires et non linéaires) en expliquant le passage de la première équation de sélection du traitement (T) à la deuxième équation de l'estimation du critère de jugement (Y) par la méthode des variables instrumentales.

Étape 3 : exploration de la pertinence de l'analyse par VI

Une description succincte des méthodes permettant d'explorer la pertinence de l'analyse par VI est proposée dans l'annexe1.

2. Typologie des variables instrumentales

Les applications de la méthode des variables instrumentales dans le domaine de la santé portent sur l'évaluation de diverses techniques d'interventions médicales, et sur l'impact de

[2] Pour obtenir ce résultat, l'instrument doit affecter le traitement de façon uniforme (propriété de monotonicité: il ne faut pas qu'une variation de l'instrument permette d'induire un accès au traitement pour certains et son contraire pour d'autres).

la consommation des substances licites et/ou illicites sur la qualité de vie des individus. Ce travail se focalise sur la littérature de l'efficacité (ou la sécurité) comparative des traitements en vie réelle (i.e. études post-inscription).

Après avoir procédé à l'interrogation de plusieurs bases de données médicales (Medline, Psych.info, Econlit...) et économiques entre 1961 et 2009, Chen et Briesacher ont réalisé une revue systématique de la littérature. 785 publications ont été identifiées. 28 d'entre elles ont été retenues dont 16 avaient des dates de publication postérieures à 2007 [15]. Pour être inclus dans la revue, les travaux retenus devaient avoir explicitement vérifié que les 3 hypothèses sur lesquelles reposait la méthode des variables instrumentales avaient été empiriquement vérifiées. Ils ont distingué ainsi 5 catégories majeures de variables instrumentales : les variations des pratiques médicales régionales et des structures de soin (13 études), les préférences des médecins (8 études), les caractéristiques des patients telles que leur statut financier et bilan médical passé (3 études) et le temps calendaire (4 études). Le tableau I décrit des exemples se rapportant à ces 5 catégories d'instruments.

Tableau I : Exemples de catégories, dimensions et type d'instruments

Catégorie	Dimension	Type de variables	Exemple d'étude
Accessibilité du service de soins	Géographique	Distance par rapport au domicile du patient	Étude : Effet du cathétérisme cardiaque sur la mortalité [7] Traitement: Cathétérisme cardiaque (oui/non) Variable instrumentale : Hôpital le plus proche du domicile du patient
Variation des pratiques ou interventions	Régionale : Offreurs de soins Établissement médecins	Fréquence régionale de l'utilisation d'un traitement (ou intervention)	Étude : Efficacité d'un traitement anti androgénique versus traitement conservateur en cas de cancer localisé de la prostate chez des hommes âgés d'au moins 66 ans [37] Traitement : Traitement anti-androgénique vs. traitement conservateur Variable instrumentale : Fréquence des traitements anti-androgénique par région
Préférences de l'offreur de soins	Médecin	Préférences du médecin (dernière prescription du médecin d'un traitement)	Étude : Effet des AINS anti-COX2 sélectifs sur les complications gastro-intestinales [9,10] Traitement : AINS anti-COX2 sélectifs vs. AINS non sélectifs Variable instrumentale: Prescription d'AINS du dernier patient reçu par le médecin
Temps de prescription (ou d'admission pour une intervention)	Temporelle : Jour, mois, année...	1. Jour de la semaine d'admission à l'hôpital 2. Date relative à une décision réglementaire sur la prescription d'un traitement médical	Étude : Efficacité de l'oxaliplatin vs. 5-FU dans la réductionde la mortalité [38] Traitement : oxaliplatin vs. 5-FU Variable instrumentale : Périodes avant et après la date de l'approbation par la FDA approbation d'oxaliplatin
Caractéristiques des patients	Individuelle-patient Caractéristiques du patient (dépenses de santé…)	Coût de traitement à la charge du patient	Étude : Effet de l'adhérence au traitement sur l'hospitalisation des patients souffrant d'insuffisance cardiaque chronique [39] Traitement : Inhibiteurs de l'Enzyme de Conversion de l'Angiotensine vs. des bétabloquants Variable instrumentale : Coût de traitement à la charge du patient (oui/non)

3. Limites de l'analyse par la VI

Les conditions d'utilisation des variables instrumentales ne sont pas toujours faciles à mettre en œuvre. Aucune des variables citées dans le tableau 1 n'est un instrument parfait. Le choix d'un bon instrument et sa justification théorique ne sont pas toujours facile. Premièrement, il n'est pas aisé de trouver un instrument très fortement associé au traitement (ou à l'exposition). Plus l'intensité de l'association entre l'instrument et le critère de jugement est faible, plus la variance de l'estimateur par la méthode VI est grande et par conséquent l'estimation de l'effet traitement est non précise. Dans ce cas, il est déconseillé d'utiliser l'approche variable instrumentale. Pire encore, si l'on dispose d'un seul instrument, qui n'est pas suffisamment lié au traitement, la stratégie consistant à créer plusieurs instruments à partir des variables disponibles dans la base de données n'est pas sans danger car non seulement elle ajoute peu d'informations mais elle aboutit à des estimations biaisées [20,21]. Ionescu-Innu et al. ont testé l'impact du choix de la variable instrumentale sur les estimations obtenues [17] en utilisant neuf instruments différents. La différence entre les résultats obtenus était grande et les intervalles de confiance différaient énormément. Le choix de la variable instrumentale doit se faire de façon prudente afin de sélectionner la variable la plus fortement associée au traitement.

Deuxièmement, même si on réussit à trouver un instrument très lié au traitement, la condition d'exclusion n'est pas empiriquement testable ; en effet, il est difficile de vérifier que l'effet de l'instrument sur le critère de jugement est exclusivement transmis via le traitement. Si l'instrument n'est pas valide (corrélé avec des caractéristiques inobservées liées au critère de jugement), l'estimateur par la VI est biaisé et comparable à celui de la méthode conventionnelle (e.g. les moindres carrés ordinaires). Dans l'exemple portant sur l'effet des AINS-COX2 sélectifs sur les complications gastro-intestinales (tableau 1), il n'est pas complètement exclu que la prescription du médecin n'affecte pas indirectement le critère de jugement (saignements gastro-intestinaux) via des traitements concomitants (cf. section application pour les études post-inscription).

Troisièmement, il est difficile de montrer que l'instrument n'est pas lié à certaines caractéristiques des patients. Dans le tableau I, l'hôpital le plus proche du domicile du patient peut être associé à certaines caractéristiques du patient (niveau socio-économique, soins habituels…) ou à d'autres co-interventions non mentionnées dans l'étude. De même, le jour de la semaine d'admission à l'hôpital peut affecter la mortalité si la prise en charge des patients admis à la fin de la semaine est très différente de celle réalisée au cours de la semaine (du lundi au vendredi).

Quatrièmement, l'analyse par la méthode VI requiert l'existence de grandes bases de données. Son utilisation sur des échantillons de petite taille est déconseillée [22] dans la mesure où elle produit des estimations moins efficientes que la méthode des moindres carrés ordinaires (MCO) par exemple. Selon les simulations de Crown et al [23] même si l'instrument est suffisamment lié au traitement et considéré comme valide, le nombre de sujets doit être assez élevé pour que la variance de l'estimation variable instrumentale ne soit pas trop grande.

La difficulté associée au choix de la variable instrumentale et la vérification des conditions de sa validité renforcent la réflexion sur la nécessité d'utiliser différentes variantes des variables instrumentales, et d'autres méthodes alternatives de contrôle des variables nommées non observables en association avec celle-ci.

APPLICATION POUR LES ÉTUDES DE POST-INSCRIPTION

Nous proposons une application de l'analyse par la méthode de la variable instrumentale sur l'innocuité comparative des anti-inflam-

matoires non stéroïdiens (AINS) en vie réelle en nous fondant sur les données utilisées par Brookhart et al [9], les travaux de Brookhart et al [9,10,16] ainsi que ceux de Rassen et al [18] et Davies et al [19]. Le choix de cet exemple est essentiellement motivé par des raisons pédagogiques : il nous permettra d'introduire et d'expliquer le principe de la méthode sans formalisation économétrique excessive, de montrer son intérêt pour appréhender la sécurité comparative des traitements dans la vraie vie, d'enrichir notamment la réflexion sur la possibilité d'exploiter des bases de données administratives françaises pour produire des évaluations des interventions médicales (utilisant par exemple la méthode de la variable instrumentale et/ou d'autres techniques existantes). Enfin les avantages et les inconvénients de la méthode des variables instrumentales pourront être présentés à la lumière de ces enseignements.

1. Analyse de l'innocuité comparative des AINS utilisés en vie réelle

Brookhart et al. [9,10] ont étudié les effets à court terme des AINS sur les complications gastro-intestinales à partir d'une cohorte de 49 919 patients âgés de 65 ans ou plus (Medicare-Pennsylvanie-États-Unis) ayant eu au moins une prescription d'AINS (sélectifs ou non) entre janvier 1999 et juillet 2002.

Les traitements comparés étaient les AINS COX2 sélectifs (coxibs) et les AINS non sélectifs. Le critère de jugement retenu reposait sur la recherche d'une différence à 60 jours entre le risque absolu des saignements gastro-intestinaux dans chacun des deux groupes. Le plan d'analyse statistique était le suivant : i) analyses univariées du risque absolu des SG avec et sans variable instrumentale. La différence de risque absolu (DRA) était approximée par la différence de proportions des saignements gastro-intestinaux entre les coxibs et les AINS non sélectifs ; ii) Les analyses multivariées (sans variable instrumentale) ont été estimées à l'aide de la procédure GENMOD (SAS version 9.1) et les modèles avec variable instrumentale ont été estimés à l'aide de la procédure IVREG (STATA version 7.0) selon la variante linéaire de la méthode VI (cf. annexe 1). Les covariables introduites dans ces modèles étaient les caractéristiques des patients (démographiques et cliniques) à l'inclusion.

Les colonnes 2 et 3 du tableau II récapitulent la répartition des variables présentes dans la base de données entre les coxibs et les AINS non sélectifs. La comparaison des variables : score de Charlson, utilisation des anticoagulants, nombre d'hospitalisations, nombre de traitements pour les complications gastro-intestinales renforce l'hypothèse de l'existence d'un biais de confusion par indication[3], et montre que les AINS COX2 sélectifs ont été prescrits aux patients ayant un risque de saignements gastro intestinaux.

Selon ces données, les AINS Cox 2 sélectifs ne semblaient pas avoir un effet protecteur sur les sujets âgés de 65 ans ou plus. Ceci peut être expliqué par une mauvaise observance, la co-prescription d'autres médicaments n'entraînant pas de complications gastro-intestinales et la présence de facteurs de risque inobservables. Les facteurs de risque des saignements gastro-intestinaux non mesurés dans les bases de données « *Health Care Claims* » sont, par exemple, la consommation de tabac, l'obésité, la consommation d'alcool, la consommation d'aspirine et les plaintes des patients au sujet des évènements indésirables

[3] Le biais d'indication est un type de biais de sélection. Dans le cas des études de post-inscription comparant deux traitements A et B, il survient quand la prescription du traitement A est liée au risque de la maladie étudiée. Ceci génère une gravité initiale non comparable entre les traitements A et B.

(par exemple, les évènements liés aux complications gastro-intestinales).

Selon le schéma d'une étude observationnelle avec banque de données médico-administratives et présence de variables non observables, la variable instrumentale est une technique potentielle pour pallier le biais des variables non mesurées contrairement à la régression standard et les variantes de l'approche des scores de propension (appariement, ajustement sur le score en présence de covariables, ou analyse stratifiée sur le score de propension).

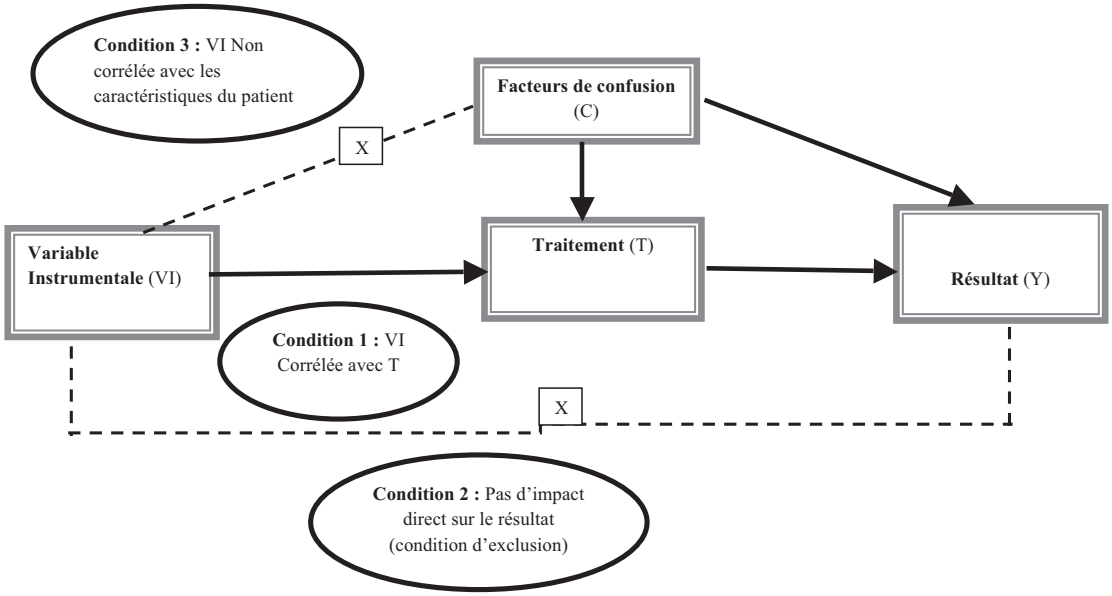
L'instrument choisi dans les études [9,10] est fondé sur la préférence de prescription des médecins. En effet, la dernière prescription d'un AINS à un patient influence largement la prescription actuelle d'un AINS. Solomon et al [24] ont montré que la préférence de prescription des médecins était un déterminant important pour la prescription des AINS COX 2 sélectifs.

Le tableau III montre que si la prescription du dernier patient reçu par le médecin est un coxib, 77,3% des prescriptions actuelles des AINS sont des coxibs. De même, si la prescription du dernier patient reçu par le généraliste est un AINS non sélectif, 54,5% des prescriptions actuelles sont des AINS non sélectifs. Ces données montrent que l'instrument (la dernière prescription d'un AINS) est bien associé au traitement (prescription actuelle), et permettent de vérifier la première condition du choix de la variable instrumentale (association entre la variable instrumentale et le traitement, représentée par une flèche entre Z et T dans la figure 2).

Tableau II : Caractéristiques des patients selon la prescription des coxibs et des AINS (avec ou sans variable instrumentale) - Source : Brookhart et al. (2006)

	COXIBS (%)		**AINS non sélectifs** (%)	
Variables	**Sans variable instrumentale**	**Avec variable instrumentale**	**Sans variable instrumentale**	**Avec variable instrumentale**
Sexe	86	84	81	84
Age > 75	75	73	65	72
Score de Charlson >1	76	75	71	73
Hospitalisation	31	29	26	27
Utilisation des anticoagulants (Warfarine)	13	12	7	10
Antécédents d'ulcère peptique	4	3	2	3
Antécédents de saignement gastro-intestinal	2	1	1	1
Traitements concomitants des saignements	5	5	4	5
Traitements pour les complications gastro-intestinales	27	25	20	24
Polyarthrite rhumatoïde	5	4	3	4
Arthrose	49	45	33	41

Figure 2 : Principe d'une Variable Instrumentale



Les colonnes 3 et 4 du tableau II montrent que les distributions des caractéristiques des patients selon la variable instrumentale (dernière prescription des AINS par le médecin) sont comparables. Ceci laisse présager que la préférence des médecins ne semble pas être liée aux caractéristiques des patients (absence de lien entre Z et C dans la figure 2). Les habitudes de prescriptions des médecins ont surement une influence forte sur le choix du traitement, mais il est peu vraisemblable que ces habitudes soient positivement liées aux déterminants de l'état de santé des patients.

Le tableau IV donne les détails de calcul de la différence du risque absolu des complications GI (selon la prescription actuelle ou conditionnelle au traitement du dernier patient reçu par le médecin). Seules les prescriptions des généralistes ont été considérées dans cet exemple.

L'effet estimé selon la méthode des variables instrumentales est donné par la formule de Wald :

$$= \frac{\begin{array}{l}[\%\ \text{des GI} \mid \text{AINS COX2 dernière prescription}] - \\ [\%\ \text{des GI} \mid \text{AINS non sélectif dernières prescription}]\end{array}}{\begin{array}{l}[\%\ \text{des COX–2} \mid \text{COX2 dernière prescription}] - \\ [\%\ \text{COX–2} \mid \text{AINS non sélectif dernières prescription}]\end{array}}$$

$$= -0.21/(77{,}3 - 54.5) = 0{,}21/22.8 : -0.92\ \%$$

Tableau III : Corrélation entre prescription actuelle d'AINS et prescription d'AINS du dernier patient reçu par le médecin – Source : Brookhart et al. (2006)

Prescription d'AINS du dernier patient reçu par le médecin (VI)	**Prescription actuelle des AINS**	
	AINS anti Cox-2 sélectifs (T=1) %	AINS non sélectifs (T=0) %
AINS anti COX-2 sélectifs Z=1	77,3	22,7
AINS non sélectifs Z=0	55,5	54,5

Tableau IV : Différence de risque absolu des saignements gastro-intestinaux selon la prescription actuelle d'AINS et la prescription d'AINS du dernier patient reçu par le médecin – Source : Brookhart et al. (2006)

	Prescription actuelle					**Prescription selon le traitement du dernier patient reçu par le médecin**				
	AINS antiCOX-2		AINS non sélectifs		DR/100 (*)	AINS antiCOX-2		AINS non sélectifs		DR/100 (*)
	N	Events	N	Events		N	Events	N	Events	
GI dans 60 jours	32 273	211	17646	110	0,03 %	25363	148	12479	99	-0,21%

(*) : Différence de risque absolu (exprimée en %).

Contrairement à l'approche de la régression conventionnelle (MCO sans variable instrumentale), l'effet estimé par la méthode des variables instrumentales montre une DRA négative et statistiquement significative au seuil de 5% (-0,92 [-1,74,-0,10] en faveur des AINS Cox2. Selon les auteurs, les résultats sans la variable instrumentale conforte l'hypothèse de l'existence d'un biais d'indication, alors que les résultats avec la variable instrumentale semblent atténuer ce biais et montrer une légère réduction des saignements gastro-intestinaux avec les AINS Cox2 en comparaison des AINS non sélectifs.

2. Discussion

Les résultats de Brookhart et al. [9] convergent avec ceux d'un essai clinique randomisé étudiant l'effet à court terme des coxibs sur les complications gastro-intestinales à 30 et 120 jours sur des patients souffrant de polyarthrite rhumatoïde [25]. Toutefois, ces résultats n'ont pas pu être comparés avec ceux d'un second essai portant sur les patients souffrant de polyarthrite rhumatoïde ou d'arthrose [26]. L'effet protecteur des COX2 n'a pas été confirmé par l'étude britannique récente qui portait sur le même sujet. En utilisant le registre britannique *« Clinical Practice Research Datalink »* (CPRD) qui couvre environ 10% de la population britannique, Davies et al [19], ont constitué une cohorte de 62 933 patients âgés de 60 ou plus, ayant consommé pour la première fois des coxibs ou des AINS non sélectifs. En utilisant les préférences des médecins comme variable instrumentale (dernière prescription d'un AINS) et dans le cadre d'une durée de traitement de 120 jours, ils n'ont pas trouvé de différence significative entre les coxibs et les AINS non sélectifs sur les risques de saignements gastro-intestinaux.

À partir des éléments de discussion avancés par Brookhart [16] sur ces résultats, nous retenons deux contraintes qui illustrent les limites de la méthode des variables instrumentales : la difficile vérification de la condition d'exclusion et l'hétérogénéité des effets du traitement.

Même si les auteurs vérifient les conditions de l'utilisation des variables instrumentales, et si les essais cliniques corroborent les résultats obtenus dans le cadre d'études observationnelles avec utilisation de variables instrumentales, rien ne garantit que la condition d'exclusion puisse être intégralement vérifiée. En effet, certains patients analysés qui appartiennent à la catégorie d'âge 65 ans et plus ont utilisé des traitements concomitants (i.e. Warfarine). De même, les associations, même faibles, d'une variable instrumentale avec les caracté-

ristiques du patient (âge, arthrose et hospitalisation) laissent présager la présence d'un sous-groupe homogène de patients à risque parmi la population analysée.

En ce qui concerne la modélisation économétrique, il aurait été préférable de réaliser des estimations fondées sur d'autres méthodes plutôt compatibles avec la nature dichotomique du critère de jugement comme la méthode 2SRI (« two stage residual inclusion) de Terza [27]. Rassen et al [18] ont essayé de pallier ce problème en proposant des techniques tenant compte de l'aspect binaire du critère de jugement (par exemple, le modèle de probit structurel et le modèle logistique à double étape) et permettant notamment l'estimation du risque relatif associé à la survenue des complications gastro-intestinales (via l'estimation des odds ratios dans le cadre du modèle logistique estimé à deux étapes). Leurs résultats confirment globalement ceux de Brookhart [9] et montrent que l'amplitude des intervalles de confiance à 95% des risques relatifs est plus faible que celle des intervalles de confiance à 95% du risque absolu sur les complications gastro-intestinales.

Les travaux sur l'innocuité comparative des AINS utilisés en vie réelle ont le mérite d'expliquer le choix de l'instrument et la discussion de l'hétérogénéité de l'effet estimé par la méthode VI mais n'explorent pas l'optimisation de la sélection d'un instrument parmi un ensemble d'instruments potentiels (exemple, le test de Sargan [28,19] dit « test de sur-identifications » permettant d'identifier le(s) «bons » instrument(s) parmi un ensemble d'instruments potentiels). De même ils ne discutent pas suffisamment la portée des tests économétriques permettant de justifier l'utilisation de la méthode VI (test d'Hausman [28,29]).

Dans la perspective d'élaboration d'un guide de bonnes pratiques sur l'utilisation de l'analyse par la méthode VI, l'utilisation de ces techniques fréquemment utilisées dans les applications de la méthode VI en économie reste à explorer dans le domaine de l'évaluation de la sécurité et/ou de l'efficacité comparative des interventions médicales en vie réelle.

3. Les enseignements à retenir

L'approche de la variable instrumentale contribue à enrichir le débat sur l'analyse de la sécurité des médicaments dans des conditions de vie réelle à partir des registres ou banques de données administratives. Avec le vieillissement de la population et l'augmentation du nombre de médicaments mis sur le marché, les études sur les bénéfices et les risques médicamenteux en conditions réelles d'utilisation dans la population âgée deviennent de plus en plus importantes. Néanmoins, le choix des préférences des médecins (i.e. la dernière prescription d'un AINS par le médecin) comme variable instrumentale n'est pas unanimement accepté si on pense que la préférence des médecins peut être, a priori, influencée par d'autres facteurs (campagnes publicitaires…), et que des prescriptions concomitantes interférant avec l'instrument risquent de remettre en question la condition d'exclusion.

Il est donc important de rechercher d'autres types d'instruments et de les comparer à la préférence des médecins afin d'optimiser à la fois la neutralité de l'instrument et sa forte association avec le traitement. Une première piste consiste à utiliser, par exemple, la technique des variables instrumentales proposée par Abrahamowicz et al. [30] adaptée à l'hypothèse réaliste de changement des préférences des médecins dans le temps, et qui permet, selon les simulations des auteurs d'optimiser l'association de l'instrument avec le traitement et d'augmenter la précision de l'effet estimé par l'approche par variable instrumentale.

Une deuxième piste à explorer des recherches méthodologiques pour comparer l'approche par variable instrumentale et le modèle de sélection d'Heckman [3]. Le modèle d'Heckman est approprié aux biais de sélection liés à des variables mesurées/non mesurées. Il permet tout d'abord de modéliser la

sélection du traitement en fonction des variables affectant la sélection, et ensuite, d'estimer l'effet du traitement en corrigeant le biais de sélection après avoir intégré un indice statistique (inverse du ratio de Mills) extrait de l'équation de la sélection du traitement.

Il est important de souligner que la variable instrumentale, à la différence du schéma randomisé, fournit un effet estimé sur le critère de jugement dit « local» dans la mesure où il concerne seulement la population sur laquelle l'instrument exerce un effet et non sur la population totale. Ainsi, l'effet estimé par la méthode VI n'est comparable à l'effet du traitement sur l'ensemble de la population que lorsque l'hypothèse d'homogénéité du traitement n'est pas remise en question. L'analyse par la variable instrumentale peut être pertinente si on recherche l'effet du bénéficie du traitement sur une population ciblée sur laquelle l'instrument a un effet (dans l'exemple de Brookhart et al. [9], l'effet recherché est celui des AINS sur les complications gastro-intestinales portant sur la population dont la dernière prescription du médecin est un anti-cox2). Toutefois, il est impératif que l'objectif de l'analyse par VI soit bien formulé pour éviter tout amalgame avec l'effet du traitement sur la population totale.

CONCLUSION

La méthode de la variable instrumentale est récemment utilisée dans les évaluations de l'efficacité ou la sécurité comparative des traitements dans des conditions de vie réelle. Elle permet de contrôler le biais induit par les variables non observables ou omises (sélection, confusion par indication...). Un instrument est une variable très liée au traitement, non liée aux caractéristiques de santé des patients, et liée au critère de jugement exclusivement via le traitement (condition d'exclusion). Ainsi, à la façon d'un plan d'expérience quasi-randomisé, un instrument pertinent offre une source de variation aléatoire permettant d'identifier un effet causal.

Néanmoins, la méthode de la variable instrumentale est très exigeante au niveau de la vérification de ses conditions d'utilisation. Dans la pratique, il n'est pas toujours facile de trouver un instrument très lié au traitement, et qui n'affecte le critère de jugement qu'à travers le traitement. Il est fortement conseillé d'abandonner cette méthode si l'intensité de l'association entre l'instrument et le traitement est faible.

La sécurité comparative des traitements médicaux en vie réelle est un champ récent d'application de l'approche variable instrumentale. La disponibilité de registres ou de banques de données administratives facilite amplement l'utilisation de cette approche. Même si les préférences de prescription des médecins demeurent un instrument discutable, l'application sur les AINS enrichit le débat méthodologique sur l'analyse de la sécurité des médicaments utilisés en vie réelle. À cet égard, les patients âgés, par exemple, peuvent être considérés comme étant une sous-population à risque pour laquelle des études après mise sur le marché sont souhaitables.

Le défi est d'évaluer l'importance des biais liés aux variables non observables et omises, de développer « en amont » des instruments dès la conception du schéma d'étude, et de planifier plusieurs types d'analyses de sensibilité. Ces dernières pourraient être fondées soit sur des variantes de la méthode des variables instrumentales permettant de retenir l'instrument le plus pertinent cliniquement et statistiquement, soit sur d'autres approches concurrentes de l'analyse d'impact des interventions médicales sur la sécurité comparative en vie réelle.

RÉFÉRENCES

[1] Ministère des affaires sociales, de la santé et des droits des femmes. Accord cadre du 5 décembre 2012 entre le comité économique des produits de santé (CEPS) et les entreprises du médicament (LEEM) *http://www.sante.gouv.fr/IMG/pdf/accord_cadre_du_051212.pdf*

[2] Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika 1983;*7:41-55.

[3] Heckman J. Sample selection bias as a specification error. *Econometrica* 1979;47:153-61.

[4] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996;81:444-55.

[5] Cameron AC, Trivedi PK. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.2009:883-8.

[6] French MT, Popovici I. That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research. *Health Econ* 2011;20(2):127-46.

[7] McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 1994;272(11):859-66.

[8] Greenland S. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* 2000;29:722-29.

[9] 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.

[10] Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf* 2010;19:537-54.

[11] Hadley J1, Yabroff KR, Barrett MJ, Penson DF, Saigal CS, Potosky AL. Comparative effectiveness of prostate cancer treatments: evaluating statistical adjustments for confounding in observational data. *J Natl Cancer Inst*. 2010;102(23):1780-93.

[12] Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*. 2007;297(3):278-85.

[13] Pizer SD. An intuitive review of methods for observational studies of comparative effectiveness. *Health Serv Outcomes Res Method* 2009;9:54-68.

[14] Swanson SA, Hernán MA. Commentary: How to Report Instrumental Variable Analyses. *Epidemiology* 2013;24:370-74.

[15] Chen Y, Briesacher BA. Use of instrumental variable in prescription drug research with observational data: a systematic review. *J Clin Epidemiol* 2011;64(6):687-700.

[16] Brookhart MA. Introduction to pharmacoepidemiology, University South Australia; Adelaide, September 2012. *http://www.vicbiostat.org.au/sites/default/files/docs/Brookhart_final_booklet.pdf*.

[17] Ionescu-Ittu R, Abrahamowicz M., Pilote L. Treatment effect estimates varied depending on the definition of the provider prescribing preference-based instrumental variables. *J Clin Epidemiol* 2012;65(2):155-62.

[18] Rassen J, Schneeweiss S, Glynn RJ, Murray A, Mittleman MA, Brookhart MA. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes, *Am J Epidemiol* 2009;169:273-84.

[19] Davies NM, Smith GD, Windmeijer F, Martin RM. COX-2 Selective Nonsteroidal Anti-inflammatory Drugs and Risk of Gastrointestinal Tract Complications and Myocardial Infarction: An Instrumental Variable Analysis. *Epidemiology* 2013;24(3):352-62.

[20] Givord P. *Méthodes économétriques pour l'évaluation de politiques publiques*. Document de travail 2010, INSEE. Paris.

[21] Greene WH. *Econometrics Analysis*. Upper Saddle River:Prentice Hall ;2003:65-83.

[22] Boef AG, Dekkers OM,Vandenbrouke JP, le Cessie S. Sample size importantly limits the usefuklness of instrumental variable methods, depending on instrument strength and level of confounding, *J Clin Epidemiol* 2014; pii: S0895-4356(14) 00232-7.

[23] Crown WH, Henk HJ, Vanness DJ. Some Cautions on the Use of Instrumental Variables Estimators in Outcomes Research: How Bias in Instrumental Variables Estimators is affected by Instrument Strength, Instrument Contamination,

and Sample Size. *Value in Health* 2011;14(8):1078-84.

[24] Solomon DH, Schneeweiss S, Glynn RJ, Levin R, Avorn J. Determinants of selective cyclooxygenase-2 inhibitor prescribing: are patient or physician characteristics more important? *Am J Med* 2003;115(9):715-20.

[25] Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, Day R, Ferraz MB, Hawkey CJ, Hochberg MC, Kvien TK, Schnitzer TJ; VIGOR Study Group. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis.VIGOR Study Group. *N Eng J Med* 2000;343:1520-8.

[26] Silverstein FE, Fraich G, Goldstein JL, Simon LS, Pincus T, Whelton A, Makuch R, Eisen G, Agrawal NM, Stenson WF, Burr AM, Zhao WW, Kent JD, Lefkowith JB, Verburg KM, Geis GS. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. *JAMA* 2000;284:1247-55.

[27] Tersa J., Bassu A., Rathouz PJ. Two-stage Residual Inclusion Estimation: Addressing endogeneity in health econometric modeling. *J Health Econ 2008*;27(3):531-43.

[28] Davidson R, Mackinnon JG. *Estimation and Inference in Econometrics*. Oxford: Oxford University Press. 1993:209-42.

[29] Hausman JA. Specification Tests in Econometrics. *Econometrica* 1978;46:1251-71.

[30] Abrahamowicz M. Beauchamp ME, Ionescu-Ittu R, Delaney Joseph AC, Pilote L. Reducing the variance of the prescribing preference-based instrumental variable estimates of the treatment effect. *Am J Epidemiol 2011;*174(4):494-502.

[31] Pirracchio R, Sprung C, Payen D, Chevret S. Benefits of ICU admission in critically ill patients: Whether instrumental variable methods or propensity scores should be used. *BMC Med Res Methodol* 2011;11:132.

[32] Johnston KM, Gustafson P, Levy AR, Grootendorst P. Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with application to epidemiological research. Stat Med 2008;27(9):1539-56.

[33] Martens EP, Pestman WR, De Boer A, Belitser SV, Klungel OH. Instrumental variables: Application and Limitations, *Epidemiology* 2006;17(3):260-67.

[34] Rassen JA, Brookhart MA, Glyn RJ, Mittleman MA. Schneeweiss S. Instrumental Variable II: instrumental variable application in 25 variation,the physician prescribing preference generally was strong and reduced covariante imbalance. *J Clin Epidemiol* 2009;62(12):1233-41.

[35] Angrist JD, Pischke JS. *Mostly Harmless Econometrics*. Princeton: Princeton University Press, 2009:113-47.

[36] Angrist JD. Estimation of limited dependent variable models with dummy endogenous regressors, simple strategies for empirical practice. *Journal of Business and Economics Statistics* 2001;19:2-28.

[37] Lu-Yao GL, Albertsen PC, Moore DF, Shih W, Lin Y, DiPaola RS, Yao SL. Survival following primary androgen deprivation therapy among men with localized prostate cancer. *JAMA* 2008;300:173-81.

[38] Mack CD, Brookhart MA, Glynn RJ, Stürmer T. Calendar Time as an Instrumental Variable in Non-experimental Comparative Effectiveness Research of Dynamic Therapies. Oral presentation, 28th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Barcelone, 2012. *http://www.pharmacoepi.org/meetings/28ICPE/presentations/aj_oral_presentations_key.cfm*.

[39] Cole JA, Norman H, Weatherby LB, Walker AM. Drug copayment and adherence in chronic heart failure: effect on cost and outcomes. *Pharmacotherapy* 2006;26:1157-64.

ANNEXE 1 : ESTIMATION PAR LA MÉTHODE DES VARIABLES INSTRUMENTALES

Équations

- Traitement : $T = \rightarrow_0 + \rightarrow_1 Z + \lambda_{11} C_{11} + \lambda_{12} C_{12} + ..+ \lambda_{1n} C_{1n} + \mp$ (1)
- Critère de jugement : $Y = \beta_0 + \beta_1 T + \lambda_{21} C_{21} + \lambda_{22} C_{22} + ..+ \lambda_{2m} C_{2m} + v$ (2)

Où : T est le traitement, Z est la variable instrumentale ; $C_{11 ...,} C_{1n,} C_{21} \ldots C_{2m}$ sont des covariables, et $\mp$ et v des termes d'erreurs.

L'équation 1 fournit les caractéristiques des individus bénéficiant d'un traitement dans le cadre d'une étude observationnelle. Dans l'équation (2) dite aussi équation structurelle, le traitement T a 2 composantes : la composante non liée avec l'erreur v et la composante liée avec l'erreur v. L'idée de base de la méthode des variables instrumentales consiste à neutraliser la composante du traitement T liée avec le terme d'erreur v.

Estimation

Variante linéaire

La méthode d'estimation classique des variables instrumentales appliquée généralement sur les modèles linéaires est la méthode des doubles moindres carrés ordinaires **(2MCO)**. Les 2MCO se réalisent en deux étapes.

Étape 1. Estimation de (1) par la méthode des moindres carrées ordinaires (MCO) pour obtenir une composante prédite et un terme résiduel**.**

-Valeur prédite de T :

$$\hat{\alpha} + \hat{\alpha_1} Z + \hat{\lambda}_{11} C_{11} + \hat{\lambda}_{12} C_{12} + \cdots + \hat{\lambda}_{1n} C_{1n} \quad (3)$$

-Terme résiduel : û

Étape 2. Intégration de la valeur prédite de T (**3**) dans **(2)** pour estimer son effet sur le critère de jugement Y.

$$\mathbf{Y = \gamma_0 + \gamma_{1*}} \text{ (valeur prédite de T) } \mathbf{+ \lambda_{21} C_{21} +..+ \lambda_{2m} C_{2m} + \nu} \quad (4)$$

L'estimateur VI de l'effet du traitement sur Y est $\hat{\gamma_1}$. La structure additive de la modélisation des 2MCO permet donc d'estimer seulement des mesures absolues de risque. Dans notre exemple sur les AINS, Brookhart et al [9] ont utilisé un modèle linéaire pour estimer la différence absolue de risque entre les AINS sélectifs et non sélectifs.

Variante non linéaire

Si le traitement et le critère de jugement sont deux variables binaires et si en plus, on veut obtenir des estimations de risque relatif, l'utilisation des 2MCO ne permet pas une bonne prédiction du traitement ou du critère de jugement en présence de plusieurs covariables. Il y a un risque d'obtention de valeurs prédites inférieures à 0 ou supérieures à 1.

Afin d'estimer des mesures relatives de risque (odds ratio, risque relatif…), d'autres méthodes d'analyse par variables instrumentales [11, 22, 31] peuvent être utilisées. Une méthode assez usuelle et adaptée aux modèles non linéaires a été proposée par Terza et al [27], qui est l'inclusion des résidus en deux étapes *« two stage residual inclusion»* (2SRI). Après avoir estimé l'équation (1), on prend les valeurs des résidus et on estime l'équation suivante :

$$\mathbf{Y = \omega_0 + \omega_1 T + \theta.\hat{u} + \lambda_{21} C_{21} +..+ \lambda_{2m} C_{2m} + \nu} \quad (5)$$

L'estimateur de l'effet du traitement par VI est $\hat{\theta}$. Si l'approche 2MCO est biaisée pour des modèles logistiques ou de Cox, l'approche 2SRI peut être utilisée par exemple avec des modèles logistiques en deux étapes.

Parmi les autres modélisations, nous citons à titre indicatif, les modèles de probit structu-

rel [19, 28,5] et la méthode des moments généralisés GMM [19, 28, 32].

Vérifications empiriques de l'analyse par VI

Les outils actuels utilisés dans le diagnostic de l'analyse par VI dans l'évaluation comparative de l'efficacité et la sécurité des interventions médicales ne permettent pas de vérifier complètement et d'une façon standardisée la validité des conditions d'utilisation de l'analyse par la VI. La première condition (l'association entre l'instrument et le fait de recevoir du traitement) peut être explorée lors de la première étape de l'estimation (équation 1) à l'aide d'un test de Fisher et le test r de corrélation partielle. Dans le cas simple d'une seule variable instrumentale si la statistique du test de Fisher est >10, l'instrument est considéré comme non faible [33]. Le calcul de ces statistiques exige la transformation d'une variable instrumentale dichotomique en variable continue. La deuxième condition (l'instrument n'a pas d'effet sur le critère de jugement excepté à travers le traitement) est difficile à vérifier. Il est impossible de tester l'absence d'association entre l'instrument et le terme d'erreur. Par contre, si l'on dispose de plusieurs instruments potentiels et si le nombre d'instruments est strictement supérieur au nombre de variables endogènes, il est possible d'effectuer un test appelé test de « suridentification » (Test de Sargan, [28,19]) permettant de se prononcer sur le(s) « bon(s) instrument(s) » à conserver parmi les instruments candidats. La vérification de la troisième condition (absence de lien de l'instrument avec les caractéristiques du patient) ne peut être vérifiée que sur les variables disponibles ; soit par une comparaison de la répartition des caractéristiques des patients selon la variable instrumentale, ou soit par la comparaison de la différence moyenne standardisée stratifiée par le traitement et celle stratifiée par la variable instrumentale [34].

Par ailleurs, dans la littérature économétrique, une vérification empirique très utile pour justifier si l'usage de la méthode VI est nécessaire consiste à tester l'égalité des estimateurs de la VI et la méthode standard (i.e. les MCO) en vérifiant, par exemple, dans l'équation (5) si le coefficient associé au résidu est significativement différent de 0. Si le résultat du test n'est pas significatif au seuil de 5%, l'usage de la méthode VI n'est pas nécessaire car l'écart type estimé de l'estimateur standard est plus faible que celui de la méthode VI. Cependant, même si cette procédure, connue sous le nom du test d'Hausman [28,29] est facilement utilisable et préprogrammée dans la majorité des logiciels économétriques, le rejet de l'utilisation de la méthode VI doit être prudemment interprété car il peut être attribuable à l'hétérogénéité de l'effet du traitement et non nécessairement à l'endogenéité de l'instrument [5,10, 35].

Il est à noter que même si le traitement et le critère de jugement sont deux variables dichotomiques, il est toujours pertinent d'estimer des variantes linéaires et non linéaires de l'approche des variables instrumentales et de comparer leurs résultats [36].

Remerciements

Ce travail a été présenté à la journée JGEM/SFES (l'efficacité comparative des traitements en vie réelle : comment choisir le bon plan d'étude, janvier 2013) et au séminaire du Conservatoire National des Arts et des Métiers sur les méthodes d'évaluations d'impact appliquées sur l'évaluation des interventions en santé publique (mai 2014). Nous remercions les deux rapporteurs anonymes pour leurs commentaires permettant l'amélioration du fond de l'article et de sa méthodologie. Nous remercions Christian Jannot et Gwenaël Le Teuff (Institut Gustave Roussy), Françoise Hamers, Alice Desbiolles, Gwénolé Briot et Eloïse Lucotte (HAS) pour leurs lectures des différentes versions de cet article.