



HAUTE AUTORITÉ DE SANTÉ
COMMISSION DE LA TRANSPARENCE

Mercredi 22 mai 2019

Seul l'avis de la CT fait foi - Occultations sous la responsabilité du laboratoire

AVERTISSEMENT

En application des articles L. 1451-1-1 et R. 1451-6 du Code de la santé publique, la HAS réalise un enregistrement des séances de la commission de la transparence (CT), de la Commission d'évaluation des dispositifs médicaux et des technologies de santé (CNEDIMTS) et de la Commission évaluation économique et santé publique (CEESP). Pour en faciliter la communication et la compréhension, la HAS a fait le choix de recourir à une transcription des débats par l'intermédiaire d'une société prestataire

Cette prestation associe une saisie directe des débats par sténotypie et une transcription assistée par ordinateur ainsi qu'une relecture médicale. L'objet de cette transcription est de permettre de tracer le déroulé des débats dans un souci de transparence et non de fournir une information scientifique validée. En effet, malgré le professionnalisme de cette prestation, il peut persister dans le texte final des incongruités ou des inexactitudes liées à l'usage d'un vocabulaire hautement spécialisé ou à la nature même des échanges verbaux. La HAS n'effectue aucune validation de ces documents.

La HAS rappelle que les seuls documents validés et opposables sont le procès-verbal de la séance et l'avis définitif de la Commission qui sont mis en ligne sur le site de la HAS.

Pour la publication des transcriptions, et dans un but de protection du secret industriel et commercial, certains mots peuvent avoir été occultés. Les occultations éventuelles sont de la responsabilité de l'entreprise exploitant le produit évalué.

Toute reprise d'un ou plusieurs extraits d'une transcription doit être accompagnée d'une mention en précisant la source et respecter la législation sur la publicité.

Les membres des commissions s'expriment à titre personnel dans le cadre de leur mission d'expertise. Les agents de la HAS (chefs de service, adjoints, chefs de projet) représentent l'institution et s'expriment en son nom.

La HAS rappelle que la connaissance des propos tenus en séance par les membres des commissions et les agents de la HAS ne peut en aucun cas justifier des contacts directs de quelque nature que ce soit avec ces personnes, lesquelles sont tenues à une obligation de confidentialité conformément à l'article R. 161-85 du Code de la sécurité sociale.

1. BRILIQUE – Renouvellement d’inscription

M^{me} GATTULLI pour la HAS.- Personne n’a à quitter la salle.

Nous accueillons M^{me} Dominique Costagliola. Il n’a pas été identifié de lien de nature à la placer en situation de conflit d’intérêts.

(L’expert, M^{me} Costagliola, entre en séance.)

M. LE PRÉSIDENT.- Merci Dominique de nous avoir rejoints, d’avoir fait ce travail et d’avoir rédigé ce rapport très complet et très intéressant.

Avant de te céder la parole, il est important de rappeler le contexte.

M^{me} GRANDE, pour la HAS.- Vous voyez BRILIQUE dans le cadre d’un renouvellement d’inscription et des résultats d’étude post-inscription. Deux études ont été soumises par le laboratoire à votre demande, deux études d’intérêt dont une réalisée par le SNIIRAM. Nous avons profité de cette occasion pour profiter de l’expertise de M^{me} Costagliola, pour qu’elle puisse vous éclairer à la fois sur l’intérêt de ces études en général et précisément ce que nous pouvons tirer de cette étude, qui est une étude post-inscription que l’on ne voit pas tous les jours.

M. LE PRÉSIDENT.- Grande spécialiste en France de ce type d’étude post-inscription, elle a rédigé beaucoup de papiers, écrit un bouquin... La référence !

██████████, pour la HAS.- Vous voyez BRILIQUE 90 mg dans le cadre de son premier renouvellement d’inscription dans l’indication prévention des événements athérothrombotiques chez les patients adultes ayant un syndrome coronaire aigu en association à l’aspirine.

Dans l’avis initial d’inscription qui date de décembre 2011, la Commission avait octroyé un SMR important et une ASMR IV par rapport au clopidogrel en termes d’efficacité. Dans cet avis de 2011, comme l’a rappelé Mathilde, la Commission avait souhaité disposer de données complémentaires recueillies en conditions réelles d’utilisation avec différents objectifs autant sur les événements indésirables que sur les caractéristiques des patients ou en termes d’efficacité.

Je laisse la parole au Professeur Costagliola.

M^{me} COSTAGLIOLA.- Sur l’étude réalisée à partir des données du SNDS, il y a un élément supplémentaire important par rapport à mon rapport : le critère de jugement prévu au protocole de cette étude, c’était le fait d’être hospitalisé, notamment pour syndrome coronaire aigu, donc c’est simplement un diagnostic principal dans le PMSI correspondant au code de syndrome coronaire aigu. En fait, ce qui a été analysé et qui est présenté dans l’article, ce sont

les gens hospitalisés pour syndrome coronaire aigu et passés en USIC. Il faut que ce soit codé en USIC. C'est un changement important, non prévu au protocole et décidé a posteriori par le comité de l'étude. C'est surprenant. J'ai regardé d'autres études réalisées à partir des données du PMSI dans le domaine cardiovasculaire, notamment une étude dans mon domaine où l'on comparait le pronostic après un infarctus chez des personnes vivant avec le VIH versus des personnes non infectées. Je n'ai pas vu d'exemple dans lequel on rajoutait la condition d'avoir été hospitalisé en USIC. C'est un élément non positif pour l'étude supplémentaire par rapport à ce qu'il y a dans mon rapport.

Les autres points qu'il faut souligner, c'est que dans l'étude, nous avons fait une analyse sous traitement et pas une analyse en ITT. Ce n'est pas le cas dans l'autre étude faite à partir de PGRX.

La méthodologie utilisée, c'est la méthode du score de propension, adaptée au fait que nous utilisons une base de données dite administrative dans notre jargon, c'est-à-dire une base dans laquelle nous avons des données qui correspondent à des dispensations de soins et à des diagnostics en cas d'hospitalisation dans lesquelles on n'a pas les facteurs de risque susceptibles d'intervenir comme des facteurs de confusion quand nous essayons de répondre à la question d'un traitement en situation observationnelle.

Du coup, les idées proposées dans la littérature, c'est de sélectionner les variables présentes avec l'idée qu'elles peuvent servir de proxy pour des facteurs de risque que nous n'avons pas. En prenant un exemple grossier, si nous prenons en compte BPCO et quelques autres pathologies pulmonaires, c'est une façon d'approcher le fait de savoir si les gens sont fumeurs. En les mettant dans un modèle, on mettra beaucoup de variables et on va quand même contrôler pour des facteurs de confusion que l'on ne mesure pas directement.

Une autre chose n'est pas détaillée dans mon rapport mais est un peu étonnante quand ils utilisent cette méthode : habituellement, on met dans le score de propension toutes les variables dont nous voulons contrôler en impact. Or, là, il y a des variables sur lesquelles ils ont directement apparié. Ils ont fait l'appariement des gens qui avaient été exposés au médicament par rapport au clopidogrel, genre, sexe ou âge ou le type de syndrome coronaire aigu, et en plus le score de propension dans lequel ils n'avaient pas mis ces variables. C'est bizarre. C'est inhabituel. C'est juste à noter.

Ceci dit, dans l'article de référence qui cite Schneeweiss. Schneeweiss montre que l'on fasse comme cela ou que l'on fasse le score de propension avec tout, cela ne change pas beaucoup les résultats. Cela a quand même un impact sur comment on va utiliser le score de propension ensuite. Le score de propension peut être utilisé de plusieurs façons. Plusieurs façons ont été décrites dans la littérature : l'appariement, nous pouvons aussi pondérer, nous pouvons aussi faire des modèles ajustés sur le score de propension. Actuellement, soit nous faisons l'appariement soit nous faisons la pondération.

Pourquoi est-ce une méthode intéressante ? C'est parce qu'il y a une étape qui échapperait si nous faisons simplement une régression multivariable : nous allons pouvoir faire un graphique

avec la distribution des scores de propension et on va voir s'il y a un certain recouvrement. La méthode s'applique en situation observationnelle. Il faut des personnes qui ont globalement des caractéristiques similaires et qui reçoivent soit un traitement soit l'autre.

Si les patients qui reçoivent le traitement A et les patients qui reçoivent le traitement B sont toujours différents, en situation d'observationnelle, vous ne pouvez rien dire, vous comparez des gens qui ne sont pas comparables. Vous ne pourrez rien dire de l'effet du traitement de façon causale.

De mon point de vue (et c'est pourquoi j'ai remis dans mon rapport le graphique avec la distribution des gens qui ont reçu le clopidogrel et ceux qui ont le ticagrélor), quand je vois ce graphique, ma réaction est de me dire que je suis dans une situation où je ne peux pas utiliser le score de propension pour conclure, parce que les gens sont relativement différents. Les pics ne sont pas du tout superposés. Eux, ce qu'ils ont fait, c'est d'apparier. Quand vous appariez, implicitement, vous allez faire deux groupes qui ressemblent au groupe des gens, ici qui reçoivent le ticagrélor. Vous allez répondre à la question : chez des gens qui ressemblent à ceux qui ont reçu le ticagrélor, est-ce que j'aurais mieux fait de donner le ticagrélor ou le clopidogrel. Est-ce la question d'intérêt ou la question d'intérêt est : pour un nouveau patient qui a un SCA, dois-je lui prescrire le clopidogrel ou le ticagrélor ? La façon d'utiliser le score de propension n'est pas la même selon que vous voulez répondre à l'un ou à l'autre.

Quand ils font cela (et c'est la conséquence du fait qu'il y a d'un overlap modeste entre les deux distributions), ils vont pouvoir apparier seulement 9000 des presque 14 000 personnes qui ont reçu le ticagrélor. Nous savons moins à cette question ils répondent puisqu'ils répondent à la question chez ceux qui ont eu ticagrélor et que j'ai pu apparier, donc une sous-catégorie qu'ils ne nous décrivent pas vraiment. On ne sait pas comment sont les 4000 non apparés par rapport aux 9000 apparés et s'il y a des différences sur ces caractéristiques qui nous permettraient de juger à qui je peux appliquer les résultats. C'est vraiment un problème majeur pour moi.

Malgré cette situation, même s'il n'y a pas un bon overlap, on trouve toujours quand même à peu près des gens à tous les niveaux du score de propension, que ce soit parmi ceux qui ont du clopidogrel et ceux qui ont du ticagrélor, quand nous regardons ces distributions. Nous aurions pu choisir une méthode permettant de prendre en compte l'ensemble des sujets en pondérant. Selon la méthode de pondération utilisée, nous pouvions choisir de répondre à la question soit chez ceux qui ressemblent à ceux qui ont eu du ticagrélor (nous sentons une pondération de type SMR) soit chez l'ensemble des nouveaux sujets et répondre à la question « pour le nouveau patient qui arrive, est-ce que je dois traiter avec A et B ? » en faisant une pondération par l'inverse de la probabilité d'être traité. Nous ne comprenons pas pourquoi ce n'est pas ce qu'ils ont fait. Dans le papier dont ils disent la méthodologie pour faire le score de propension de haute dimension, derrière, ils faisaient une pondération par le SMR dans les exemples où ils illustraient leur méthode. En lisant le papier sur le même site, ils pouvaient voir comment faire.

Au final, dans les analyses de sensibilité qu'ils réalisent, ils vont faire un modèle de régression multiple dans lequel ils vont pouvoir inclure l'ensemble des 13 916 patients. Ce modèle sera

ajusté sur les variables d'appariement qu'ils ont laissées à part (l'âge, le sexe, etc.) et sur le score de propension. Le fait d'ajuster dans un modèle multivarié sur le score de propension, c'est une méthode utilisée dans la littérature dans des scores de propension. Normalement, cela marche moins bien, mais bon. Au final, dans tout ce qu'ils ont fait, c'est ce que je retiens comme moins mauvaise analyse, puisque c'est une analyse qui permet d'inclure l'ensemble des 13 916 patients.

Évidemment, les résultats ne sont pas extrêmement différents. Les ordres de grandeur du risque sont relativement similaires. Mais si on est avec le couperet de la signification statistique, dans l'analyse des 13 000, sur le critère primaire de jugement, on n'atteint pas la signification, alors qu'on l'atteignait dans l'analyse restreinte aux 9000. Mais comme on ne sait pas à quoi ces 9000 correspondent, ce n'est pas l'analyse qu'on a envie de retenir.

En revanche, les résultats sur le décès sont robustes, quelle que soit la méthode utilisée, montrant un avantage pour le ticagrélor versus le clopidogrel pour les décès. C'est pour l'étude basée sur le SNDS.

Pour l'étude AREMIS, cette étude est basée sur un registre de cas de syndrome coronaire aigu recrutés dans plus de 300 centres en France. Ils ont réalisé une ITT, mais le critère de jugement est uniquement le risque de récurrence d'un infarctus, infarctus non fatal ou décès cardiaque. Je veux bien qu'il y ait un peu des biais, mais décès cardiaque, je trouve que c'est toujours compliqué d'utiliser pas seulement le fait que la personne soit décédée mais la cause de décès. C'est l'enfer de valider une cause de décès. C'est toujours délicat.

Ils font une approche de type cas témoins appariés.

Là aussi, nous constatons qu'ils ont pu appairer seulement 1047 des 1243 cas d'infarctus récurrent éligibles. Même s'il y a beaucoup de précautions dans la réalisation de l'étude, il y a forcément un plan d'étude plus susceptible d'être entaché de biais de sélection voire de facteurs de confusion, d'autant plus que tous les cas ne viennent pas du registre. Ils n'avaient pas tout à fait assez de cas dans la puissance. Ils ont inclus dans les mêmes centres des cas en plus de ceux du registre. Nous pouvons d'ailleurs nous demander ce qu'étaient ces cas survenus dans les mêmes centres mais qui n'étaient pas dans le registre. Ou cela ne s'appelle pas un registre.

Ce qui est embêtant, qui n'est pas bien discuté, est que quand on fait ce genre de chose, on a l'objectif d'arriver à une interprétation causale. Nous savons qu'il y a des problèmes. On discute les sources de biais et de facteurs de confusion point à point en donnant des éléments pour dire pourquoi on pense que cela ne joue pas de rôle. Là, c'est à peine si nous avons cela. Ils vont bien dire à un moment qu'il ne faut pas exclure qu'il y ait de la confusion résiduelle. D'ailleurs, le fait de trouver des résultats aussi différents de 1, c'est probablement dû à cela. Peut-être entièrement ! Je ne trouve pas que la discussion de cela... Comment arrivez-vous à convaincre le lecteur qu'il faut interpréter le résultat de façon causale alors que nous sommes en situation observationnelle ? Cet exercice est délicat. Je ne trouve pas que la discussion du papier rend compte à cet aspect.

Je sais qu'il y a des problèmes de classe, mais on s'arrange pour évoquer ces problèmes et les discuter. Il y a des exemples fameux, excellents, de certains des auteurs de l'article où ils ont su faire. C'est faisable.

Personnellement, au total, sur l'ensemble des deux études, je retiendrai plutôt l'étude sur les données du SNDS, plutôt l'analyse réalisée sur le modèle de Cox incluant l'ensemble des sujets, avec quand même une interrogation majeure par rapport au changement du critère de jugement puisqu'avec le critère de jugement tel que prévu au protocole, nous ne trouvons aucun bénéfice au ticagrélor.

M. LE PRÉSIDENT.- Merci beaucoup pour cette analyse fine, parfois complexe pour les non-spécialistes.

M^{me} COSTAGLIOLA.- C'est un des soucis de ces méthodes. Elles ont beaucoup de puissance, mais le diable est dans les détails. Il est difficile d'être sûr que l'on va tout bien regarder et ne pas laisser passer de soucis. Ce n'est pas strict forward, ce n'est pas presse-bouton, c'est très complexe, y compris dans la mise en œuvre.

M. LE PRÉSIDENT.- Quand tu dis « apparier » pour répondre à la question : « est-ce que ticagrélor fait mieux ? », pourquoi est-ce différent de dire : « Si j'ai un nouveau patient, qu'est-ce que je dois choisir ? »... Je comprends ta réponse.

M^{me} COSTAGLIOLA.- Les patients ticagrélor sont très différents de ceux qui sont sous clopidogrel.

M. LE PRÉSIDENT.- C'est à cause de ça ?

M^{me} COSTAGLIOLA.- Oui.

Ce n'est pas illégitime en post-AMM de demander si les gens qui ont reçu le produit, c'était à bon escient. En appariement, il faudrait pouvoir apparier plus de 90 % des sujets pour répondre à cette question. Ce n'est pas la même question que : un nouveau arrive, est-ce que je dois lui donner A ou B ?

M. LE PRÉSIDENT.- Pour le praticien, c'est un peu lié quand même. Il sera tenté de choisir celui des deux qui aura...

M^{me} COSTAGLIOLA.- C'est un élément que nous n'avons pas pris en compte, mais nous sommes dans la période où le produit arrive, donc il faut une accoutumation. Nous comprenons que l'on peut choisir les gens à qui on prescrit le nouveau traitement en fonction de ce qui a été dit ou de notre sentiment. Il n'est pas illogique que les patients ne se ressemblent pas.

M. LE PRÉSIDENT.- Le fait que les pics soient décalés, c'est lié à quoi ? C'est dû à l'appariement ?

M^{me} COSTAGLIOLA.- Non, c'est la distribution sur l'ensemble des sujets. Cela dit globalement, avec un indicateur résumé qui convient toutes les variables : les patients ne se ressemblent pas.

Ce n'est pas les mêmes à qui on a prescrit. S'il n'y a pas des gens relativement semblables à qui on a prescrit A et B, on ne peut rien dire de la différence entre A et B en situation observationnelle.

Pour moi, l'intérêt essentiel à recourir à un score de propension plutôt qu'à une régression multiple, c'est que nous voyons s'il est raisonnable ou pas de le faire. Si je regarde le graphique, ma conclusion est de dire : en situation observationnelle, je ne peux pas répondre.

M. LE PRÉSIDENT.- C'est clair.

M. Le P^r MERCIER.- Je comprends votre difficulté. Finalement, ce ne sont pas les mêmes types de patients. D'un autre côté, il y a eu récemment des mises au point, notamment dans le *Lancet* fin 2017, sur les différents essais qui étaient bout à bout. Quelque part, l'avantage de l'essai randomisé, c'est qu'a priori, nous partons d'une population bien équilibrée entre les deux bras. Cela facilite peut-être un peu plus l'interprétation.

J'en conclus que nous sommes toujours dans une balance risque/bénéfice entre les risques de rethrombose ou de thrombose ailleurs (d'autres endroits qui peuvent se thromboser) et les risques de segment. J'ai retenu que les anticoagulants (anti-thrombines K) faisaient mieux en type de thrombose et faisaient plus saigner le cerveau. Finalement, on balance entre différents antiplaquettaires plus ou moins efficaces entre clopidogrel, ticagrélor et aspirine qui commence à sentir son ancienneté. La preuve, la dernière mise au point est une controverse publiée dans le *New England* la semaine dernière. Certains sont pour et d'autres sont contre. Ce ne sont que des avis. Le débat continue.

L'avantage des essais randomisés, c'est que les patients en principe étaient bien équilibrés entre les groupes. Là, c'est la difficulté de votre exercice qui est brillant. Vous concluez que ce ne sont pas les mêmes populations et que l'on ne peut pas conclure.

M^{me} COSTAGLIOLA.- Complètement. C'est pourquoi que, pour l'enregistrement, on demande des essais randomisés et on répond à la question : au nouveau qui arrive, est-ce qu'on lui donne A ou B ?

M^{me} le D^r DEGOS.- Il nous arrive avec une certaine fréquence de demander des études post-inscription, et quand nous les avons, nous sommes généralement déçus. Y aurait-il quelque part une guideline à laquelle pourraient se conformer les gens qui font l'étude post-inscription ? Ce sont des milliers de malades que l'on a étudiés et qui permettraient d'y voir plus clair. Là, nous avons l'impression de beaucoup de travail pour un résultat qui n'est pas satisfaisant.

M^{me} COSTAGLIOLA.- Il y a deux aspects dans l'étude. Nous aurions pu être dans une situation où il y avait plus de comparabilité entre les groupes. Du coup, nous pouvions répondre à la question en situation observationnelle. C'est l'effet de l'utilisation du produit qui fait qu'on ne peut pas faire de comparaison directe en situation observationnelle. On ne peut pas savoir à l'avance. Cela dépend de la façon dont les produits vont être utilisés en arrivant sur le marché. C'est complexe. Vont-ils remplacer ? À quelle vitesse ? Est-ce seulement dans une niche ?

Après, il y a l'aspect de le faire méthodologiquement. Là, il y a deux soucis, alors que c'est une équipe très expérimentée qui connaît très bien l'utilisation des données du SNDS : c'est le côté changement de critère de jugement. C'est quand même embêtant. Puis, il est vrai qu'historiquement, nous faisons plutôt l'appariement, mais dans la littérature, nous voyons de plus en plus de la pondération qui permet d'inclure tout le monde, et nous sommes plus confortables avec cela. Ce sont des défauts de cette étude spécifique : changer de critère et ne pas prendre de méthode prenante en compte tous les sujets. Nous savons à quelle question nous répondons. Quoi qu'il en soit, l'interprétation sera complexe. Ce n'est pas exactement les mêmes patients qui ont reçu les deux produits.

Il y a eu deux aspects : comment sera utilisé le produit ? Quand on demande des études en vie réelle, c'est aussi pour savoir cela. Là, on voit que cela n'a pas été utilisé exactement chez les mêmes patients, les patients qui n'avaient pas exactement le même type de syndrome coronaire aigu.

M^{me} le D^r DEGOS. - C'est imparable.

M^{me} COSTAGLIOLA. - Oui, mais nous avons appris les caractéristiques des patients traités par A et B.

M. le P^r GUEYFFIER. - Nous pouvons débattre à l'infini de la capacité des scores de propension, des différentes méthodes pour soutenir des affirmations causales. Un des autres intérêts et aspects de ce type d'études post-inscriptions observationnelles, c'est de vérifier que nous ne nous écartons pas trop des recommandations et de l'autorisation de mise sur le marché.

Là, nous sommes dans un cas un peu particulier. Nous avons des informations sur l'intérêt de donner ce type de traitement pendant un an après un syndrome coronarien aigu. Nous avons évalué ensuite dans une étude spécifique si poursuivre au-delà d'un an amène quelque chose. Après un an suivant le syndrome coronarien aigu, nous avons un équilibre des accidents hémorragiques graves et des accidents ischémiques prévenus qui fait que le rapport bénéfice net s'annule. Cela ne vaut pas le coup de poursuivre le traitement.

Je ne sais pas si la question avait été précisément posée aux industriels d'explorer cette dimension intéressante de la poursuite ou non au-delà d'un an de l'accident qui rend les patients éligibles au traitement. C'est une question très intéressante, pas forcément évidente à étudier sur des bases de données, mais qui est quand même assez accessible et qui ne posera pas trop de problèmes d'interprétation.

Y a-t-il des données sur la durée du traitement après l'accident rendant les patients éligibles ?

[REDACTED], pour la HAS. - Si je peux me permettre, nous avons demandé à un an post-événement.

M. le D^r KOUZAN. - J'ai une question naïve. Est-ce que le graphique que vous avez mis dans votre rapport, montrant que les deux populations ne sont pas comparables ne devrait pas être

non pas un prérequis mais un critère de qualité obligatoire dans les études observationnelles de ce genre ?

M^{me} COSTAGLIOLA.- Quand on utilise le score de propension, il faut forcément avoir ce graphique d'une part et d'autre part un autre graphique qui est bien fourni par le même supplément de l'article qui est les différences standardisées entre les deux groupes, pour vérifier que toutes les différences standardisées de toutes les variables qu'on a prise en compte sont inférieures à 10 %, ce qui montre que l'on a réussi à faire un équilibre.

Pour moi, dans un article où on fait du score de propension, cette figure doit forcément être fournie. Si je ne l'ai pas et si je suis reviewer, je demande qu'on me la montre. C'est vraiment l'intérêt essentiel du score de propension. Au-delà, je suis d'accord : quand on est, dans une étude de cohorte, en situation observationnelle, c'est un graphique obligatoire. Dans l'autre étude, étude cas témoin, on ne peut pas fournir ce type de graphique. Je ne peux pas le demander. Nous ne pouvons pas calculer cela. Si c'est une étude de cohorte, il faut fournir ce graphique (c'est indispensable) et de préférence faire un modèle multivarié.

De toute façon, quand vous avez toutes les variables à bord, vous ne pourriez pas faire un modèle multivarié où vous mettriez 500 variables. Il y a intérêt à faire derrière le score de propension. Dans 90 % des cas, si vous avez toutes les variables pertinentes de facteur de confusion, les deux méthodes vont donner plus ou moins la même chose. Le seul intérêt de passer par le score de propension, c'est de voir s'il est raisonnable de continuer.

M. LE PRÉSIDENT.- Pourquoi on a encore recours à l'appariement puisque la pondération paraît préférable ?

M^{me} COSTAGLIOLA.- Historiquement, il y avait différentes choses. Au début, on faisait des analyses stratifiées par des strates de scores de propension. On a pu faire des modèles multivariés. Nous mettons le score. Dans mon rapport, je cite un article intéressant comparant toutes ces méthodes. Il explique bien, dans une situation exceptionnelle, comment cela peut donner des choses extraordinairement différentes et pourquoi. C'est une survivance, à mon avis. Personnellement, j'ai plutôt tendance à faire de la pondération, mais je ne me souviens plus si au début, je n'ai pas fait un peu d'appariement. Mais je n'en fais plus.

M. LE PRÉSIDENT.- Faute avouée est à moitié pardonnée. C'est une évolution qui paraît inéluctable et qui n'est pas plus compliquée que l'appariement.

M^{me} COSTAGLIOLA.- Non, c'est plus simple. C'est dans les logiciels. Ils permettent de faire l'appariement. C'est trop facile. Vous écrivez moins d'une ligne de commande, cela le fait automatiquement, vous ne contrôlez... Et cela fait peur, parce que c'est complexe et il faut vérifier que l'on n'a pas laissé de choses dans un coin. Sur le fait de faire des analyses pondérées et d'avoir le poids dans une variable, tous les logiciels permettent de le faire depuis longtemps. Ce n'est pas très difficile.

M. le D^r BLONDON.- Compte tenu de la significativité discutable du critère principal, est-ce qu'on peut interpréter le critère secondaire, qui est la survie, qui est quand même un critère d'intérêt majeur, qui semble très en faveur du ticagrélor ?

M^{me} COSTAGLIOLA.- Je ne me souviens plus ce que cela donne quand ils font l'analyse avec leur critère primaire sur ce critère. J'ai regardé les résultats du composite, mais sur le critère qui n'est pas composite, je ne sais plus ce que l'on trouve. Sur le critère primaire initial, c'est en faveur de l'essai quoi qu'il arrive. C'est la seule chose de robuste. Je l'ai souligné dans le rapport. C'est aussi un critère qui compte. Comme là c'est décès toutes causes, il est relativement indiscutable.

M. Le P^r THIERRY.- J'ai eu le temps de vérifier dans le papier que les chercheurs ont travaillé en toute indépendance par rapport à l'industriel. C'est noté dans le papier. Une fois que l'on a dit cela, je suis quand même surpris, parce qu'il y a l'identification d'un sous-groupe qui vous étonne, celui qui passe par les USIC. On dit souvent que c'est pour avoir un résultat positif. C'est fréquent dans les publications. C'est moins la mode de publier des résultats négatifs. C'est étonnant. Vous remarquez qu'ils n'utilisent pas la bonne technique. Comment expliquer ces deux biais dans la mesure où ils étaient parfaitement indépendants, qu'il n'y avait pas de clause de confidentialité, et qu'ils disent que l'industriel n'a pas participé au traitement des données, n'a pas participé à la rédaction, etc. ?

M^{me} COSTAGLIOLA.- Je ne sais pas. Je ne veux pas lancer des hypothèses. C'est un laboratoire très compétent et très connu. Maintenant, je n'ai pas forcément compris l'argument avancé pour avoir changé le critère de jugement principal. Ce qui m'étonne, c'est qu'ils le changent a posteriori de leur protocole alors que c'est une équipe qui a l'habitude de travailler sur les données du SNDS dans le domaine cardiologique.

M. Le P^r THIERRY.- C'est dans les papiers très influencés par l'industrie que l'on trouve ce genre de méthodes d'habitude.

M^{me} le D^r DEGOS.- N'y a-t-il pas eu plus de place en USIC pour mettre les infarctus ? Des pratiques médicales peuvent varier en fonction du contexte.

M^{me} COSTAGLIOLA.- La proportion des infarctus passés à l'USIC n'était pas la même dans les deux groupes. Après, c'est un autre point général des données du SNDS. Il ne faut pas oublier que c'est un codage qui est fait par les hôpitaux pour gagner de l'argent. Donc, il n'a pas un objectif médical, il a un objectif de coût. Cela peut influencer le codage. Je ne vois pas comment ce serait confidentiel. Cet effet est connu. Il faut y penser quand on interprète des données issues du SNDS. Il faut savoir que ces données sont collectées dans un but économique.

M^{me} le D^r DEGOS.- Sont-elles collectées de façon simultanée ? Est-ce que cela ne s'est pas décalé dans le temps ?

M^{me} COSTAGLIOLA.- Il n'y a pas de remarque sur cela. Je ne saurais pas dire. Mais cela m'étonne que cela puisse être différentiel. Pourquoi aurions-nous plus cela en diagnostic

principal dans un séjour non USIC pour des gens qui reçoivent certains médicaments par rapport à d'autre ? Je n'ai pas bien compris l'argument. Je ne vois pas comment cela pourrait être différentiel. Je pense que nous essayons toujours de faire un codage favorable à l'hôpital. Quand nous faisons cela, nous ne regardons pas que les gens ont le médicament A ou B.

M^{me} le D^r DEGOS.- Je pose la question du temps parce qu'il y a eu une appropriation par les cliniciens du codage et une amélioration sensible...

M^{me} COSTAGLIOLA.- Nous sommes à une période où la qualité du codage est bonne. L'étude est réalisée à une période récente. Je n'ai plus les dates en tête. Le problème de la qualité du codage peut s'expliquer au début on a mis le PMSI dans les hôpitaux d'étude mais pas sur la période d'étude couverte ici.

M. LE PRÉSIDENT.- Pour nous, les études post-inscription sont importantes surtout avec des développements parfois compliqués. Ce type de résultats, il y a des avantages, il y a des publications de très bon niveau, il y a des limites. La conclusion qui est très prudente consiste à dire que cela ne remet pas en cause la situation du ticagrélor versus clopidogrel.

M^{me} COSTAGLIOLA.- Ce n'est pas le sens de ma conclusion.

M. LE PRÉSIDENT.- C'est plus prudent. Je suis d'accord.

M^{me} COSTAGLIOLA.- Je ne suis pas membre de cette commission, c'est à vous de faire la conclusion. J'ai dit ce que je retiendrai (dans tous les chiffres, ceux que je retiens) avec un élément de bémol, non écrit dans mon rapport initial, qui est que de toute façon ces résultats sont dans le cadre d'une situation où j'ai changé mon critère principal de jugement par rapport au protocole. Après, j'avais bien différencié, comme vous l'avez souligné, le résultat sur le critère événement et le résultat sur le critère mortalité. Nous avons vérifié ensemble que sur le critère mortalité, quelle que soit l'analyse regardée, nous avons des résultats robustes. Après, vous décidez ce que vous faites avec ces résultats. C'est en ce sens que la conclusion n'était pas dans un sens ou l'autre.

M. LE PRÉSIDENT.- La conclusion orale est plus précise que la conclusion écrite.

Quelqu'un d'autre veut-il intervenir ?

(Réponse négative)

Merci beaucoup pour cette belle étude et cette belle analyse.

(L'expert, M^{me} Costagliola, quitte la séance.)

Avez-vous des remarques complémentaires ou des questions notamment de la part des méthodologistes ?

M. le D^r KOUZAN.- Après avoir entendu ce genre d'exposé, je me demande si les études post-inscription ne devraient pas être limitées à juste la recherche d'effets secondaires, d'une meilleure évaluation de la iatrogénie. Pour le reste, nous nous perdons dans la méthodologie et elles deviennent non conclusives parce que ce ne sont pas des études randomisées.

M. LE PRÉSIDENT.- C'est très intéressant que la robustesse des résultats pour aboutir à d'autres conclusions en termes d'efficacité prête à caution.

M. le P^r CLANET.- C'est quand même un sujet majeur que ces études de propension. Dans ma spécialité particulièrement, au cours des 10 dernières années, il y a un nombre considérable d'études qui ont hiérarchisé des effets thérapeutiques sur des approches identiques et dans lesquelles il y a toujours, effectivement, une critique méthodologique. Quand nous regardons ces études de façon approfondie avec des gens qui ont l'habitude, nous avons toujours un biais qui arrive, qui, finalement, remet un peu en question les conclusions de ces études. C'est toujours un peu difficile.

Nous voyons ici. On nous présente deux études : l'une sur laquelle nous avons vu combien elle était très prudente sur la façon de l'interpréter en disant que finalement, les patients ne sont pas les mêmes ; la deuxième étude sur laquelle elle dit que ce n'est pas interprétable et ce n'est pas bon. Nous sommes dans une situation d'incertitude dans laquelle nous avons quand même un petit élément de robustesse (j'aimerais que vous me le confirmiez) sur la mortalité, ce qui est quand même, dans des études observationnelles, un élément important. Maintenant, est-ce que ce problème de différence de mortalité est exclusivement lié à l'effet du médicament ? C'est une question. C'est compliqué.

M. le P^r GUEYFFIER.- Je crois qu'effectivement, il faut accorder aux études observationnelles leur place qui est la leur et reconnaître qu'elles n'ont pas d'aptitude à maîtriser les biais comme l'essai randomisé. Sinon, nous n'aurions pas besoin de faire essai randomisé. Si nous pouvions faire de l'observationnel avec les bases Sécurité Sociale, nous aurions une puissance telle et si nous pouvions tirer les mêmes conclusions que d'essais randomisés, nous serions ravis.

Effectivement, il y a des limites intrinsèques qu'il faut savoir reconnaître. Je me permets de faire une petite publicité pour un article que nous avons récemment commis dans *Therapy* avec Michel Cucherat en identifiant les différents biais de ce type d'études observationnelles.

Je reviens sur le graphique montrant la répartition des scores de propension dans les deux groupes ticagrélol et clopidogrel. C'est assez logique. C'est le but d'utiliser toutes les variables disponibles pour savoir si les patients qui ont eu du ticagrélol sont différents des autres. En utilisant ce score, on range les patients selon la propension, la probabilité d'être traités par clopidogrel. D'un côté, nous avons plus de patients sous ticagrélol. Heureusement ! Cela veut dire que le score a une certaine efficacité. À gauche, nous trouvons les patients qui ont eu plus de clopidogrel. Il est logique d'avoir cette distribution, un certain recouvrement, et Dominique Costagliola a dit qu'il en fallait un minimum de recouvrement pour tirer des conclusions. Sinon, on ne pouvait pas faire d'analyse tranche par tranche de score de propension. En même temps, nous n'avons jamais de dissociation comme elle l'évoquait. Là, je trouve que le niveau de

recouvrement permet de faire des analyses qui ont la capacité des scores de propension, c'est-à-dire d'ajuster sur ce qui est disponible. Maintenant, ce qui est disponible, ce n'est pas l'ensemble des faits sur lesquels s'appuient les médecins pour prescrire un traitement et non pas un autre, que l'on ne pourra pas capturer comme le fait le hasard à travers des essais randomisés.

M^{me} GRANDE, pour la HAS.- Sur les essais cliniques et données en vie clinique, pour rappeler la doctrine, vous écrivez qu'il n'y a pas lieu d'opposer ces deux types d'études (essais cliniques randomisés, études observationnelles). Vous rappelez que l'étude clinique randomisée reste le schéma de référence pour démontrer l'efficacité d'un médicament. Cela signifie que les études observationnelles ne peuvent pas se substituer aux essais cliniques, néanmoins elles apportent des compléments d'information importants en termes d'usage, d'efficacité dans la vraie vie. C'est votre positionnement, qui correspond à vos discussions d'aujourd'hui.

M. LE PRÉSIDENT.- On est un peu plus positif sur les études observationnelles. On dit qu'elles sont toujours analysées au vu de l'ensemble des données d'efficacité disponible sur le traitement et peuvent être un des éléments conduisant la HAS à valoriser ou maintenir une ASMR.

M^{me} GRANDE, pour la HAS.- C'est l'un des éléments de l'appréciation, mais cela ne se substitue pas un essai clinique.

M. LE PRÉSIDENT.- C'est une discussion importante. On demande des études post-inscription pour compléter notre interprétation du produit, notamment en termes de tolérance et d'usage. À chaque fois que nous parlons d'efficacité, nous disons que franchement, nous avons du mal à conclure. Tu le dis aussi sur l'évaluation différente de l'efficacité. C'est lourd à mettre en œuvre, une EPI. Les laboratoires le reprochent. Ils disent que nous demandons des EPI, puis nous en soulignons les limites. Il faut insister sur le fait qu'elles ont de la valeur à nos yeux, mais on ne peut pas leur demander plus qu'elles ne peuvent donner.

M. Le P^r MERCIER.- Je répondais dans le même sens. La phase IV, souvent parce qu'il y a beaucoup plus de malades, permet éventuellement de faire émerger un signal. Sur les essais randomisés, en cardiologie, nous sommes dans les 10 000 ou 15 000 patients ; en pédiatrie, on est plutôt en dessous de 1000 patients. L'étude observationnelle permet d'essayer de faire émerger un signal qui n'est pas apparu au niveau de la phase III. C'est un complément. D'un autre côté, cela coûte tellement cher qu'il ne faut pas donner à ces études observationnelles le complément de substitution qui pourrait nous être reproché parce que les laboratoires pourraient dire : « Contentons-nous de faire une étude observationnelle. Cela coûte trop cher de faire des essais randomisés. » Et on aura de moins en moins de médicaments reposant sur des essais randomisés. Nous en avons fortement débattu récemment.

Je suis frappé dans la spécialité de pédiatrie que quelque part, les médecins urgentistes pédiatres aux urgences ont réussi à faire émerger un grand nombre de problèmes quand ils ont travaillé en réseau ensemble. Je voudrais citer ce que l'on appelle PECARN, c'est-à-dire Pediatric emergency care ambulatory research network. Le fait de mettre 25, 30 ou 40 permettait d'avoir

des bases de données suffisamment importantes pour poser les données d'un problème qui permettraient ensuite de donner le contour raisonnable en termes d'estimation de puissance, etc. de faire des essais randomisés. En fait, c'est complémentaire. Il y a cet aspect observationnel tenant compte de l'aspect collaboratif de pouvoir poser les problèmes ou les données pour un essai randomisé. Vous avez eu l'expérience de certains essais randomisés qui nous a soumis au niveau de la pédiatrie, au niveau de la mucoviscidose... À chaque fois, nous avons des difficultés pour avoir le nombre de patients nécessaire pour répondre aux questions. Mais le déblaiement par ces bases de données et de travail collaboratif ont permis de poser les données. Après, je crois qu'il faudrait se limiter dans nos demandes d'études post-observationnelle ou post-inscription parce qu'on ne peut pas leur faire lire ce qu'elles ne peuvent pas dire.

M. le D^r ROSENHEIM.- Je pense que quand la Commission demande des études post-inscription, elle ne devrait demander, sauf exception concernant essentiellement la toxicité, que des études analytiques reposant notamment sur des analyses avec score de propension.

De façon plus générale, je voudrais revenir sur un point. Les essais randomisés ont une grande valeur parce qu'ils ont une bonne validité interne, car ils ont une randomisation, ce qui permet d'ajuster sur tous les facteurs de confusion connus et inconnus, ce qui n'est pas possible dans des études observationnelles. En revanche, les études randomisées ont un défaut : elles ont habituellement une très mauvaise validité externe. C'est difficile d'extrapoler leurs données à la population cible. Les patients sont sélectionnés. Ils sont volontaires. Il y a des critères d'exclusion très nombreux qui généralement ne sont pas repris dans leur totalité dans le RCP. Donc des patients qui n'auraient pas pu être inclus dans un essai randomisé vont recevoir malgré tout le traitement pour ces raisons. En revanche, les études observationnelles, qui sont critiquables du point de vue de la validité interne, puisqu'elles souffrent de biais, ont en revanche une validité externe excellente. Les patients reçoivent réellement le traitement. Et c'est la population qui reçoit le traitement qui est analysée.

Je pense qu'il serait temps d'arrêter d'opposer les deux types d'études, que ces deux types d'études qui sont complémentaires : bonne validité externe pour les études observationnelles et bonne validité interne pour les essais randomisés. Une fois de plus, je pense que les études présentées comme étant des études post-inscriptions, c'est-à-dire des études uniquement descriptives, ne nous apportent rien du point de vue information. Il faudrait arrêter d'en demander (c'est gaspillage de temps pour tout le monde et d'argent pour les firmes) et ne demander que des études analytiques.

Il y a une lettre dans le *JAMA cardiology* pouvant faire penser que le ticagrélor a une activité antibactérienne, notamment sur les Gram positifs y compris sur les staphylocoques résistants à la méticilline. Cela a été constaté parce que les patients qui avaient une pneumonie et recevaient du ticagrélor avaient un meilleur pronostic que ceux qui n'en recevaient pas.

M. LE PRÉSIDENT.- Je ne pense pas qu'on ait opposé ces deux types d'études, mais tu précises des choses concernant les études post-inscription que nous demandons en étant plus précis que nous ne le sommes. Les études observationnelles, pour le bon usage, cela en fait partie...

M. le D^r ROSENHEIM.- Est-ce que le bon usage est dans nos missions ? C'est dans les missions de l'assurance maladie. Je ne suis pas sûr que ce soit dans les missions de la Commission de la Transparence.

M. LE PRÉSIDENT.- Si, aussi !

Par ailleurs, tu abordes le fait de ne demander que les études avec score de propension. C'est valable pour les études analytiques que nous demanderions, en précisant quand nous voulons des études analytiques, donc avec score de propension. Mais le score de propension est quasi inéluctable pour toi ?

M. le D^r ROSENHEIM.- Il y a d'autres méthodes, mais la méthode actuellement utilisée, c'est le score de propension. Les autres ont un peu disparu. Il y avait une méthode de randomisation, pseudo-randomisation consistant à choisir les patients habitant à la même distance du centre où ils étaient traités. Les résultats n'étaient pas inintéressants. C'est une façon de randomiser les gens.

M. LE PRÉSIDENT.- On peut randomiser sur de nombreux paramètres. C'est sûr.

M. Le P^r THIERRY.- La validité externe des études observationnelles parle aux patients. C'est souvent qualifié d'étude en vie réelle. La vie réelle, c'est très important pour les patients. Je veux dire par là que les études post-inscription devraient se multiplier parce que le niveau de preuve au moment de l'AMM est moins exigeant que par le passé au niveau réglementaire, à commencer par la FDA et par l'EMA.

Je voudrais rebondir. Christophe dit que c'est cher. Non, ce qui est cher, c'est de laisser surutiliser sur le marché (ou sous-utiliser) un médicament qui en vie réelle, après une évaluation initiale avec un niveau de preuve faible, pourra entraîner des dépenses importantes. De la même façon que d'aller jusqu'au critère principal de survie globale ajoute un an et est très important dans les études qui s'arrêtent éventuellement trop tôt, il va devenir indispensable de faire des études en vie réelle observationnelles pour vérifier dans les conditions de prescription avec patients représentatifs (cela dépasse le cadre ici) l'efficacité d'une prescription. Si on ne le fait pas, les dépenses induites sont énormes. Comparée aux dépenses induites justifiées, la dépense d'une étude observationnelle est très faible.

M. LE PRÉSIDENT.- Cela complète ce que nous disions. Je crois que c'était une discussion intéressante. Dans la mesure où c'est un renouvellement d'inscription, nous devons nous prononcer à nouveau sur le SMR. Il était important. Qui est pour maintenir ?

(Il est procédé au vote.)

Mercredi 22 mai 2019

17

SMR important : unanimité.

Nous n'avons pas à voter sur l'ASMR.

Seul l'avis de la CT fait foi - Occultations sous la responsabilité du laboratoire

