



TRANSCRIPTION DES DÉBATS

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.

Avis économique XENPOZYME (SANOFI-AVENTIS) - ECO-EFFI 572 - Adoption

M^{me} PARIS.- Nous passons à l'avis économique de XENPOZYME que vont nous présenter nos chefs de projet. On a déjà vu ce dossier en groupe technique, donc j'espère qu'on pourra aller assez rapidement. Nous avons trois personnes qui ont travaillé sur ce dossier et qui nous le présentent. Allez-y.

Un chef de projet du SEM.- Bonjour à tous. Nous allons vous présenter XENPOZYME. Avant de débiter, nous souhaitons remercier encore une fois nos rapporteurs, Claire Granon et Lionel Perrier, pour tous les échanges et leur support. On va vous présenter XENPOZYME dans le traitement enzymatique substitutif des manifestations non neurologiques, du déficit en sphingomyélinase acide de type B et A/B chez les patients pédiatriques et adultes.

Il s'agit d'une primo-inscription qui a obtenu le statut de médicament orphelin le 19 septembre 2001, une AMM centralisée le 24 juin 2022, dans la même indication que celle demandée au remboursement, à savoir le traitement enzymatique substitutif des manifestations non neurologiques du déficit en sphingomyélinase acide de type B et A/B chez les patients pédiatriques et adultes. Il faut savoir que XENPOZYME était prescrit sous ATU nominative depuis avril 2020, et en accès précoce pré-AMM depuis le 23 mai 2022.

La maladie ASMD consiste en un déficit de l'enzyme sphingomyélinase, ce qui entraîne une accumulation de la sphingomyéline avec une atteinte multi-organes qui peut être neurologique et/ou viscérale.

La maladie concerne moins de 90 patients en France, elle est considérée donc comme très rare, prématurée et pouvant entraîner le décès.

Il existe trois types présentés ici, du moins sévère au plus sévère. On a :

- le type B qui peut débiter à tout âge et n'a pas vraiment d'atteinte neurologique ;
- le type A/B, qui débute dans l'enfance et comporte une atteinte neurologique et viscérale avec un pronostic qui est moins favorable que le type B ;
- enfin le type A, qui est la forme la plus sévère, mais qui n'est pas concernée par l'AMM.

Il s'agit d'un traitement enzymatique substitutif qui permet de réduire les manifestations cliniques, notamment en améliorant la fonction pulmonaire et en réduisant le volume de la rate et du foie. Aucun traitement n'est disponible en France à ce jour.

L'industriel revendique une indication superposable à l'AMM, un SMR important et un ASMR II au prix revendiqué de ■■■ euros par milligramme. Aucun impact sur l'organisation des soins n'est revendiqué. Comme évoqué tout à l'heure, la population cible est de 90 patients.

La HAS a fait la démarche auprès du service des usagers pour obtenir la contribution de deux associations de patients, à savoir l'Association française Vaincre les maladies lysosomales et l'Association espagnole des patients atteints d'ASMD. Celles-ci mettent en avant notamment l'impact de la maladie sur la qualité de vie, l'aspect invalidant dû à l'attente multi-systémique et l'amélioration de la qualité de vie constatée par les patients qui ont été inclus dans l'essai clinique.

Concernant les données d'efficacité, la source des données qui ont été mobilisées, nous avons plusieurs essais cliniques, et une étude que je vais détailler. En jaune, vous avez les essais cliniques qui incluent les patients pédiatriques et en bleu ceux qui incluent les patients adultes.

On a l'étude ASCEND-Peds, qui est une étude mono-bras qui concerne une population pédiatrique. Ensuite, nous avons l'étude DFI qui est aussi une phase I, aussi mono-bras et qui concerne cinq patients adultes. Ces deux études sont ensuite regroupées pour une étude LTS, étude de long terme qui va regrouper à la fois des patients adultes et des patients pédiatriques.

Enfin, nous avons l'étude ASCEND aussi appelée ETS, qui est la seule étude comparative, mais sur une durée de seulement 52 semaines. Parce que dans sa phase d'extension ASCEND, les patients du bras placebo vont rejoindre le groupe traité par olipudase alfa. ASCEND extension est une étude mono-bras. Il faut bien faire la distinction entre la phase d'extension de ASCEND et l'étude long terme LTS.

Enfin, nous avons l'étude SPHINGO-100, qui est une étude prospective d'histoire de la maladie sur une durée de onze ans. Elle a été agrégée avec les essais pour agrandir la cohorte de patients.

Deux analyses ont été déposées, une pour la population pédiatrique et une pour la population adulte. L'objectif ici est conforme. Cependant, l'horizon temporel vie entière fait l'objet d'une réserve mineure en raison de l'incertitude générée par l'absence de données comparatives, notamment pour la population pédiatrique dans l'étude ASCEND-Peds, et la population adulte au-delà d'un an.

Il s'agit d'un modèle de Markov à neuf états de santé, avec en violet clair les trois niveaux de gravité de splénomégalie et en violet foncé trois niveaux de gravité de l'altération de la fonction pulmonaire. Le patient peut transiter entre les états de volume splénique et les états DLCO de

manière indépendante, et la transition d'un état extrême à un autre sans passage par un état intermédiaire n'est pas autorisée. L'industriel justifie l'utilisation d'un modèle de Markov par le caractère chronique et évolutif de la pathologie et l'absence de données de mortalité robustes recueillies dans les essais du fait de l'espérance de vie, qui est relativement longue chez les patients.

Cependant, le nombre d'états de santé soulève des limites, notamment concernant leur alimentation dans un contexte de maladie rare, avec peu de données. Je laisse maintenant la parole pour vous présenter la modélisation de l'efficacité.

Un chef de projet du SEM.- Merci. On va aborder les points spécifiques à la modélisation de l'efficacité. D'abord les sources de données. Pour tenir compte du maximum d'informations, toutes les données cliniques et l'étude de l'histoire naturelle ont été agrégées. Ils ont fait un pool de données pour estimer les matrices de transition. Cela nous a semblé acceptable dans un souci d'apport du maximum d'informations. Également, l'industriel a tout de même discuté de la potentielle hétérogénéité introduite par l'agrégation de différentes études, à la fois avec des designs différents, mais aussi des critères définis parfois différemment entre les sources.

On va se concentrer maintenant sur la méthode qui a été appliquée pour estimer les probabilités de transition. Chez l'enfant comme chez l'adulte, les probabilités de transition ont été estimées à partir d'une comparaison naïve entre les résultats des essais cliniques, et ceux de la cohorte historique SPHINGO-100, conduisant à une estimation non robuste de l'efficacité d'olipudase alfa.

Quand on parle d'estimations non robustes, c'est qu'il n'y a aucune technique d'ajustement statistique qui a été mise en œuvre par l'industriel pour tenter de rendre comparables les populations comparées et ainsi reproduire les conditions d'un essai randomisé.

En lieu et place de l'ajustement statistique que nous aurions attendu, l'industriel a uniquement réalisé une analyse quantitative des différences entre les deux groupes comparés et a tenté de discuter l'impact éventuel de ces différences en termes d'impact, d'efficacité et sur les résultats de la modélisation.

On vous a représenté les deux tableaux que l'industriel nous a présentés, où vous voyez à la fois chez l'adulte et l'enfant, un certain nombre de variables qui sont statistiquement différentes entre les deux groupes comparés. En particulier, il y a des différences sur des critères d'intérêt de la pathologie, comme la sphingomyéline par exemple, ou également l'âge moyen chez l'enfant.

Ces différences sont susceptibles de biaiser l'estimation de l'effet traitement issu de la comparaison naïve. L'incertitude autour des résultats de la comparaison naïve qui a été utilisée en analyse de référence n'est pas explorable.

On va maintenant différencier un peu plus les justifications de l'industriel chez l'adulte et chez l'enfant, du pourquoi ils ont utilisé la comparaison naïve. Chez l'adulte, c'est très clair qu'au vu de la documentation qui nous a été apportée, on considère que l'approche retenue n'est pas justifiée compte tenu du peu d'éléments documentant et discutant ce choix. Il n'est donc pas possible pour nous de savoir si d'autres approches plus robustes auraient pu être envisagées, puisque l'industriel ne discute pas du fait qu'ils l'ont ou qu'ils auraient pu tester des techniques d'ajustement.

L'industriel a fourni une analyse en scénario utilisant uniquement les données comparatives de l'essai ASCEND. On obtient un RDCR réduit de 28 %. Toutefois, ce scénario est interprété avec précaution puisque, et ce sont les arguments même de l'industriel, l'extrapolation des données du bras *standard of care* repose sur un suivi d'un an uniquement, puisqu'il y avait un switch qui était autorisé dans ASCEND. Par ailleurs, l'industriel n'a pas fourni l'exploration complète de l'incertitude avec les analyses de sensibilité déterministes et probabilistes, ce qui ne nous permet pas de retenir cette analyse.

On propose une réserve majeure sur la méthode d'estimation des probabilités de transition, avec la formulation suivante : la comparaison naïve entre les données cliniques agrégées évaluant l'olipudase alfa et l'étude historique SPHINGO-100, conduisant à une estimation non robuste de l'efficacité d'olipudase alfa qui amplifie l'incertitude associée aux résultats du modèle économique et qui n'est pas conforme aux recommandations de la HAS.

On va maintenant s'attarder aux cas pédiatriques. De la même manière, une comparaison naïve a été réalisée. Cette fois-ci, on a un peu plus de documentation puisque l'industriel nous dit qu'il a testé deux méthodes de comparaison indirecte : une méthode que l'on pourrait appeler par ajustement et une méthode un peu plus sophistiquée que l'on retrouve souvent dans les dossiers par score de propension.

Toutefois, on se rend compte qu'avec ces deux méthodes, on arrive après ajustement, avec une réduction extrêmement importante de l'échantillon après ajustement, que ce soit avec les deux méthodes, mais c'est encore plus amplifié avec la méthode de score de propension. C'est la dernière colonne, vous voyez que pour le bras *standard of care*, l'effectif est réduit de quasi 80 %.

Ce qu'on a regretté dans le dossier, c'était le manque de discussion de la méthode par score de propension et d'explications aussi de pourquoi les effectifs étaient si réduits après ajustement. On a fini par comprendre que c'était, d'une part, lié aux nombreuses données manquantes dans

les sources de données mobilisées, et également au fait qu'après ajustement, les populations étaient tellement différentes qu'il y a eu une étape supplémentaire de tronquage des patients avec des poids élevés. Cela montre qu'il y avait une forte hétérogénéité dans les groupes comparés de base.

Finalement, l'industriel a privilégié les résultats de la comparaison naïve, puisque les deux méthodes testées ne permettaient pas de renseigner les probabilités de transition des neuf états de santé du fait de la réduction importante de l'effectif après ajustement. Toutefois, et on le redit, la comparaison naïve n'est pas acceptable puisque cela conduit à une estimation non robuste. C'est notamment le cas pour la DLCO où on se rend compte quand les résultats de la comparaison indirecte qui sont présentés dans la note d'intérêt thérapeutique manquent avec une comparaison indirecte. L'effet sur la DLCO était statistiquement non significatif.

Ce résultat, qui n'a pas été discuté, questionne finalement de la validité de l'approche retenue de modéliser un effet traitement sur la DLCO à partir des résultats de la comparaison naïve. De la même manière, on propose une réserve majeure sous le même intitulé que celle que j'ai citée précédemment.

Au final, avec la comparaison naïve, ils estiment plusieurs matrices de transition. Ils en estiment quatre matrices sur des périodes différentes pour le bras olipudase alfa et trois matrices pour le bras *standard of care*. Là, ce sont les matrices chez l'enfant sur le critère DLCO. Vous avez également deux matrices de ce type pour le critère volume de la rate, chez l'enfant et chez l'adulte.

Vis-à-vis de l'hypothèse d'extrapolation de l'effet traitement, vous voyez dans le tableau de gauche, les matrices de transition de l'année 2 à 5 ans sont les mêmes que de l'année 5 à la fin de la simulation. L'industriel suppose que l'effet traitement observé entre 2 et 5 ans dans les études à long terme est maintenu au-delà de 5 ans.

On a questionné cette hypothèse de maintien de l'efficacité au cours du temps avec les experts externes, et ils nous ont confirmé que c'était une hypothèse acceptable et plausible, du fait notamment du mécanisme d'action de l'olipudase. Mais dans cette pathologie, il semblerait qu'il n'y ait pas de production d'anticorps énorme.

Par ailleurs, l'industriel a présenté tout de même une analyse de sensibilité testant une stabilisation de l'effet traitement après 5 ans et on observe une variation marginale du RDCR. C'est pourquoi on n'a pas proposé de réserve sur ce sujet.

On va passer ensuite aux événements intercurrents. Le premier point concerne les événements indésirables. On observe une hétérogénéité entre les sources de données mobilisées pour

l'efficacité et la tolérance, qui n'a pas été justifiée, avec un impact inconnu sur les résultats, mais qui est attendu faible en raison de la part marginale des coûts liés aux événements indésirables sur le coût total. À la différence de l'efficacité où il y avait un certain nombre d'études qui ont été agrégées, pour la tolérance, les événements indésirables ne viennent que de ASCEND et ASCEND-Peds, les études principales. On propose sur ce point une réserve mineure sur les sources de données mobilisées pour estimer la tolérance.

On propose également une réserve importante sur l'estimation du taux d'observance en maintenance chez les adultes, qui est uniquement basée sur les données ASCEND Extension. Or on se rend compte que dans l'étude LTS qui a aussi inclus des adultes, et qui a été mobilisée pour estimer l'efficacité, l'observance était de 96,8 %. La conséquence, c'est qu'il y a probablement un biais associé au lien entre l'estimation des coûts d'acquisition et l'estimation de l'efficacité, avec une baisse du coût d'acquisition sans impacter l'efficacité, ce qui est susceptible de sous-estimer le RDCR.

La réserve est importante, puisque l'impact de l'observance sur le RDCR est fort. On a formulé la réserve de la façon suivante : incertitude entourant l'estimation du taux d'observance chez l'adulte en phase d'entretien, et ce d'autant plus que les données d'efficacité appliquées correspondent à une meilleure observance.

Enfin, le dernier point vise l'intégration des complications. On se retrouve dans un cas où, faute de données, effectivement, ils ont eu le recours à plusieurs sources de données. Les taux annuels de complications sont issus de l'étude SPHINGO-300 et les risques relatifs sont issus soit de ASCEND, soit de SPHINGO-100, en fonction du type de complications.

On propose une première réserve sur le manque de discussion sur la comparabilité des sources utilisées pour intégrer les complications et l'absence de discussion aussi sur la plausibilité clinique, d'intégrer une probabilité de survenue de complications identiques chez l'adulte et chez l'enfant.

On propose aussi une réserve mineure sur l'application d'un effet traitement sur la réduction des complications hépatiques, fondées sur l'étude ASCEND et sa phase d'extension, et qui est maintenue au-delà des données cliniques disponibles, ce qui ne nous semble pas acceptable dans le contexte de comparaison naïve des données d'efficacité.

Toutefois, la réserve n'est que mineure pour cette seconde réserve, puisque l'absence d'intégration de cet effet traitement sur les complications hépatiques a été testée en analyse de sensibilité et l'impact est faible.

Je passe pour la suite.

Un chef de projet du SEM.- Merci. En ce qui concerne la validité, nous avons jugé cela acceptable. À noter qu'il n'y a aucun exercice de validation externe qui a pu être mené, faute de données disponibles. Toutefois, ceci est compréhensible dans le contexte de maladies rares.

Je vais revenir maintenant sur l'estimation des scores d'utilité. Comme nous l'avions présentée lors du GT Eco, la source des données retenue en analyse de référence pour estimer les scores d'utilité est une étude de vignette qui a été développée spécifiquement pour la modélisation afin de renseigner les neuf états de santé du modèle.

L'étude de vignettes se décompose en deux phases.

- Une phase I, la phase pilote qui a reposé sur 22 participants, patients cliniciens et aidants, afin d'ajuster la description des états de santé qui fondent les vignettes,
- une deuxième phase, la phase de valorisation, reposant sur 200 participants représentatifs de la population générale anglaise, qui a permis de valoriser les scores d'utilité par la méthode du time trade-off. Les 200 participants étaient divisés en deux groupes, évaluant successivement les vignettes en population adulte et pédiatrique de façon séquentielle, pour tenir compte de l'effet d'ancrage.

Vous avez ici des exemples de vignettes, par exemple des vignettes adultes avec la vignette parfait état de santé en rose, en haut ; la vignette M en verte, qui représente une atteinte légère en termes de DLCO, sans atteintes en termes de volume splénique et de foie ; et vous avez la vignette K en bleu qui représente un état sévère d'ASMD. Vous voyez que, mis à part la vignette parfait état de santé, les vignettes sont décomposées en trois catégories : les atteintes respiratoires sur la DLCO, les atteintes sur le foie et la rate, et les autres variables de qualité de vie, que ce soit sur les activités quotidiennes, monter des escaliers, aller au travail, les infections, les hospitalisations, l'apparence et l'aspect émotionnel.

Ici, vous avez des vignettes enfants pour la population pédiatrique. Nous vous avons remis la vignette grise d'informations de contexte, puisqu'il faut savoir que dans l'étude de vignettes, il y a toujours une vignette de contexte qui permet vraiment de se replacer dans la peau, ici, d'un enfant de huit ans.

Comme je l'ai dit tout à l'heure, dans la phase II, ce sont des participants représentatifs de la population générale adulte anglaise. Pour répondre aux vignettes, dans l'estimation des scores d'utilité pédiatrique, les participants ont dû se mettre dans la peau d'un enfant de huit ans atteint d'ASMD, avec les différents niveaux d'atteinte. Par exemple, ici, une atteinte faible en termes de DLCO et une atteinte moyenne en termes de rate et de foie pour la vignette orange Q, et un état

sévère ASMD pour la vignette K en bleu. On retrouve encore les trois catégories, comme chez les adultes.

Nous avons mis ici les scores d'utilité introduits dans le modèle. On voit que dans la population adulte, ces scores varient de 0,918 pour le meilleur état de santé à 0,33. pour l'état de santé le plus dégradé. Dans la population pédiatrique, ces scores varient de 0,939 pour le meilleur état de santé jusqu'à 0,450 pour l'état de santé le plus dégradé.

Quels sont les arguments de l'industriel pour privilégier l'étude de vignettes par rapport au questionnaire de qualité de vie générique collecté dans l'essai ASCEND et ASCEND-Peds ? C'est tout d'abord le manque de sensibilité et d'exhaustivité des dimensions génériques de ces questionnaires, dans le cadre de la pathologie qui nous concerne ici, l'ASMD, notamment sur la détérioration des capacités respiratoires, les infections pulmonaires, l'impact de la splénomégalie, etc. Ces questions génériques auraient conduit à des résultats contre-intuitifs avec des scores d'utilité pour les états de santé les plus dégradés qui pouvaient être supérieurs aux scores d'utilité pour des états de santé moins dégradés.

L'industriel justifie aussi le recours aux vignettes par un manque de robustesse statistique en raison du faible nombre d'observations dans les essais. Cela est encore plus marqué dans la population pédiatrique, comme nous le verrons sur la *slide* suivante. Il justifie enfin le recours à l'étude de vignettes par les effets de seuil sur l'auto-évaluation des patients et l'adaptabilité des patients dans un contexte de pathologie chronique. C'est un phénomène bien connu dans la littérature pour les pathologies chroniques.

Toutefois, cette méthodologie présente de nombreuses limites et sources d'incertitudes. En effet, l'effectif de la phase pilote, la phase I, est limité et manque de documentation sur la répartition et les caractéristiques des 22 participants.

De plus, la complexité de la pathologie de l'ASMD et ses atteintes biologiques rendent difficile l'appropriation par des participants qui ne sont pas familiers et qui n'ont pas expérimenté cette pathologie. Cela est encore plus marqué pour l'évaluation par des participants adultes non familiers de la pathologie de scores d'utilité pédiatrique.

À noter aussi que la valorisation a été faite auprès d'un échantillon représentatif de la population générale anglaise. La transposabilité est donc incertaine au contexte français, que ce soit en termes de préférence, et aussi en lien avec le système de soins, les différences entre le système français et le système britannique.

De plus, on a une incertitude quant à la prise en compte des complications. En effet, il y a un risque de double compte de leur effet, avec un risque de double compte des désutilités qui

pourraient être appliquées directement *via* les descriptions des vignettes et qui sont appliquées par ailleurs *via* des utilités, nous reviendrons dessus juste après. Cela n'a pas été discuté par l'industriel.

Enfin, il y a une absence de méthodologie standardisée pour la réalisation de vignettes qui peut introduire des biais de réponse en fonction de la description des vignettes qui est réalisée dans l'étude.

Nous vous avons remis la recommandation du guide de la Haute Autorité de Santé vis-à-vis de l'étude de vignettes. Le guide dit : « L'estimation des scores d'utilité par une approche fondée sur la révélation des préférences pour un état de santé fictif *via* des vignettes ou par une échelle visuelle analogique n'est pas acceptée, même dans le cadre d'une analyse de sensibilité. » Ici, dans ce dossier, l'étude de vignettes est en analyse de référence.

Nous avons remis ici les scores d'utilité collectés dans les essais, puisqu'il y a tout de même eu des questionnaires génériques qui ont été administrés lors des essais. Pour les adultes, c'était le questionnaire EQ-5D-5L administré à l'inclusion à 26 semaines et à 52 semaines. On voit que le nombre d'observations disponibles sont tout de même assez restreints, notamment pour les états de santé très dégradés ou l'état de santé le plus favorable. Les scores estimés à partir du modèle mixte mesures répétées, qui sont introduites dans le deuxième tableau, sont en effet très proches d'un état de santé à un autre.

Pour l'enfant, pour la population pédiatrique, c'est le questionnaire pédiatrique *pediatric quality of life inventory*, Peds-QL, qui a été administré dans l'essai clinique ASCEND-Peds, et les réponses étaient faites soit directement par les patients, soit par les parents.

Comme je le disais tout à l'heure, l'un des arguments du laboratoire pour rejeter ces questionnaires est de dire qu'il y a un manque de robustesse dû au faible nombre d'observations. On voit ici qu'en ce qui concerne Peds-QL à l'inclusion, nous avons 10 observations chez les patients enfants et 14 chez les parents, avec des scores qui allaient de 0,86 à 0,81 respectivement. À 52 semaines, il y avait 9 observations chez les patients enfants et 11 chez les parents, avec des scores allant de 0,92 à 0,90 respectivement.

Une analyse en scénario a été réalisée par l'industriel à partir des données EQ-5D issues de l'essai ASCEND, chez les adultes. Ces données ont été appliquées dans les deux sous-populations, à savoir la population adulte et la population pédiatrique.

On observe que le RDCR augmente de plus de 45 % chez l'adulte, et plus de 65 % dans la population pédiatrique. Ainsi, l'impact est important et favorable aux produits évalués. Le fait de retenir une étude de vignettes semble favorable aux produits évalués au vu des résultats de cette

analyse en scénario. Ainsi, l'incertitude est forte concernant l'estimation des scores d'utilité dans le contexte spécifique de la pathologie de l'ASMD, dans laquelle l'évaluation de la qualité de vie est pourtant un élément central.

Nous proposons donc une réserve majeure avec le libellé suivant : « La méthodologie utilisée pour mesurer et valoriser les états de santé en termes de score d'utilité à partir d'une étude de vignettes présente de nombreuses limites contribuant à créer une forte incertitude sur les résultats en QALY et n'est pas conforme au guide méthodologique en vigueur pour l'évaluation économique à la HAS. »

Comme je le disais tout à l'heure, nous avons une incertitude au niveau des complications et de l'estimation des désutilités liées aux complications. Celle-ci est doublée d'une incertitude sur les sources permettant d'évaluer ces désutilités. Nous proposons une réserve mineure avec le libellé suivant : En l'absence de donnée de désutilités spécifiques aux complications d'intérêt et associées à l'olipudase alfa dans l'ASMD, la transposabilité à la population simulée des sources mobilisées dans d'autres pathologies demeure incertaine. L'impact des désutilités sur le RDCR est faible. Puisqu'en effet, pour estimer ces désutilités liées aux complications, ils ont retenu différentes sources issues de la littérature, évaluant des pathologies proxys de la complication évaluée.

En termes de coûts, les postes de coûts considérés sont classiques : coût d'acquisition, coût d'administration, coût des événements indésirables liés au traitement, coût de transport et coût de suivi en pratique courante. Nous proposons une réserve mineure sur la population pédiatrique, puisque l'estimation des coûts d'acquisition d'olipudase alfa chez l'enfance soulève des incertitudes liées à la représentativité du poids moyen à l'entrée de la cohorte de 23,32 kilos estimé par à partir du z-score. Il y a une incertitude également sur l'évolution du poids des patients traités par olipudase alfa dans le temps. L'impact est toutefois limité aux huit premières années de modélisation, puisque l'âge moyen de la population pédiatrique est de 11 ans.

En termes de résultats. Pour rappel, les résultats sont invalidés puisque la CEESP propose deux réserves méthodologiques majeures. Sur un horizon temporel vie entière, le RDCR de l'olipudase alfa revendiqué par l'industriel était de 1 942 000 euros par QALY dans la population adulte et 1 660 000 euros par QALY dans la population pédiatrique.

L'exploration de l'incertitude a montré que le RDCR était notamment fortement influencé dans la population adulte par le poids et l'âge adulte, ainsi que l'observance, comme l'a rappelé le chef de projet tout à l'heure. Dans la population pédiatrique également, ce sont les mêmes variables, à savoir le poids, l'âge adulte et l'observance à l'âge adulte qui influe le plus sur le RDCR.

Nous vous avons remis ici une synthèse des réserves proposées, à savoir une réserve mineure sur l'horizon temporel, la réserve majeure sur la comparaison naïve dans les deux populations, les réserves mineures sur les événements intercurrents, événements désirables, complications et observance, avec la réserve importante sur l'observance, la réserve majeure sur la méthodologie de l'étude de vignettes utilisée pour mesurer les scores d'utilité associés aux états de santé, une réserve mineure sur la désutilité liée aux complications, une réserve mineure sur les coûts dans la population pédiatrique uniquement et une réserve mineure sur les analyses de sensibilité.

Un chef de projet du SEM.- Nous allons vous projeter la conclusion qui reprend le contenu et le détail des deux réserves majeures qui invalident l'ensemble des résultats soumis à la CEESP chez l'adulte et l'enfant.

On a rajouté un élément à la fin de ce paragraphe qui souligne que dans le contexte d'une pathologie rare avec peu d'observations dans les essais, il n'est pas à exclure que le développement d'un modèle simplifié, avec le regroupement d'états de santé les plus pertinents cliniquement, aurait pu permettre de pallier à certaines limites liées à l'estimation des probabilités de transition et des scores d'utilité.

Par ailleurs, au regard des résultats des analyses de sensibilité, on rappelle à l'intention du lecteur qu'il est important de noter que l'observance au traitement, le poids des patients ainsi que les pratiques de prescription, notamment la possibilité d'utilisation des reliquats, sont des variables qui affectent le plus l'olipudase alfa. Pour les patients traités à l'hôpital, la mise à disposition prochaine d'un conditionnement d'un flacon de 4 milligrammes pourrait permettre de minimiser les pertes de reliquat à l'avenir.

Ensuite, depuis la version qu'on a pu vous envoyer, on a rajouté un paragraphe sur l'impact budgétaire. Dans le cadre de ce dossier, le chiffre d'affaires était estimé à [REDACTED] d'euros. L'industriel n'a donc pas présenté d'impact budgétaire. Toutefois, dans l'avis CT, on a une estimation de la population susceptible d'être traitée supérieure de 10 % par rapport à celle que l'industriel revendique. On peut se poser la question de la pertinence, de l'éligibilité à un impact budgétaire avec cette population cible retenue par la commission Transparence.

On a juste souhaité souligner que compte tenu de l'incertitude sur l'effectif de la population qui sera effectivement traitée par olipudase alfa, ce montant pourrait être plus élevé, en parlant du montant du chiffre d'affaires. Une analyse d'impact budgétaire aurait été utile pour documenter l'impact financier attendu pour l'Assurance maladie de l'introduction d'olipudase sur le marché français.

Pour la conclusion de la commission, on propose tout d'abord une phrase qui dit : l'efficience du produit n'est pas démontrée en raison des deux réserves méthodologiques majeures qui invalident les résultats de l'évaluation économique dans les deux populations. D'une part l'estimation des probabilités de transition fondée sur une comparaison naïve ne permet pas d'estimer de manière robuste l'efficacité entre l'olipudase alfa et les soins standards. D'autre part l'estimation et la valorisation des scores d'utilité à partir d'une étude de vignettes qui présente des biais est à l'origine d'une forte incertitude sur les résultats en QALY, et ce d'autant plus que l'estimation de la qualité de vie des patients atteints d'ASMD est un élément central.

Enfin, on a proposé une formulation pour insister sur le fait que dans le contexte de maladie rare telle que l'ASMD, la CEESP souligne la nécessité de mener en amont une réflexion sur le design des essais cliniques en veillant à y inclure le maximum de patients en vue de recueillir le plus de données exploitables pour l'évaluation économique de ces produits.

En termes de données complémentaires, on a indiqué que les résultats de l'analyse de l'efficience ont besoin d'être corroborés par les données comparatives documentant l'efficacité à long terme et la qualité de vie des patients traités par olipudase alfa.

Voilà pour le projet d'avis.

M^{me} PARIS.- Merci beaucoup pour cette présentation qui était très claire. Claire Granon et Lionel Perrier sont les rapporteurs de ce dossier. Souhaitez-vous prendre la parole ?

M^{me} GRANON.- D'abord je vous remercie beaucoup, les chefs de projet, de la présentation très complète, très claire et qui reprend vraiment toutes les discussions qu'on a eues. Je les ai un peu bombardés de questions par mail, et chaque fois, j'ai pu avoir les réponses.

Je suis un peu désolée de ce dossier parce qu'il y avait une très grande attente des patients. On a tout de même des patients pédiatriques qui sont lourdement handicapés, des patients adultes aussi. L'arrivée de ce médicament ouvrait tout de même des portes. Des patients ont été traités et ont fait des retours d'information qui étaient très positifs. Je trouve tout à fait dommage que ce dossier soit traité de cette façon au niveau de l'efficience. Il y a d'une part les problèmes de comparaison naïve et surtout cette histoire de vignette qui m'a beaucoup troublée. On aurait pu faire beaucoup mieux, déjà dès le design de l'essai, mais ensuite, peut-être dans les explications complémentaires qui sont un peu insuffisantes. On a l'impression qu'on a fait le produit et qu'on s'est dit que ça passerait parce qu'il y avait tout de même l'enthousiasme des patients, etc.

Le dossier n'est pas suffisant, et je le regrette. Je souscris complètement aux conclusions du travail qui a été fait.

M^{me} PARIS.- Merci beaucoup. Lionel.

M. PERRIER.- Beaucoup de choses ont été dites. Je voudrais remercier les chefs de projet, ainsi que Claire, pour tout le travail réalisé sur ce dossier. Il s'agit d'une maladie rare, avec un nombre très limité de patients, la population cible étant estimée à 90 patients. Le dossier a fait l'objet d'un échange technique dense, extrêmement dense. Malgré l'amélioration du dossier, un certain nombre de points méthodologiques et posent encore problème. Cela a été largement décrit au cours de la présentation.

À noter que les points qui posent problème avaient été tous évoqués lors de l'échange technique. On peut citer la méthode d'estimation des probabilités de transition à partir d'une comparaison naïve. L'industriel a bien justifié son choix dans la population pédiatrique, et les résultats ont été positionnés au regard des méthodes de comparaison indirecte par ajustement et par score de propension, mais cette justification manque dans la population adulte.

Si je reviens sur la modélisation, la mise en œuvre du modèle de Markov est pertinente, mais peut-être aurait-il fallu simplifier les états de santé pour enseigner de façon plus robuste les transitions de probabilité, l'évaluation de l'utilité qui repose sur une étude de vignette en analyse de référence, avec toutes les limites encore qui ont été largement décrites dans la présentation.

Je voudrais souligner que la spécificité de cette maladie rare, conjuguée au très faible nombre de patients, complexifie l'analyse médico-économique. Nous avons longuement échangé entre les rapporteurs et les chefs de projet en charge du dossier pour savoir si nous pouvions tout de même présenter le RDCR, par exemple pour la population pédiatrique seule, ou en coût par année de vie gagnée. Mais compte tenu des nombreux biais et des non-conformités au guide méthodologique, nous n'avons pas retenu cette approche.

Je voudrais simplement terminer par la petite mention que nous avons proposée dans la conclusion, c'est-à-dire qu'il est très important et encore plus important d'anticiper, dans ces pathologies rares, la meilleure conception et la meilleure organisation possible des études cliniques. Je vous remercie.

M^{me} PARIS.- Merci beaucoup à tous les deux. C'est très clair également. Fanny.

M^{me} MONMOUSSEAU.- Bonjour à tous. J'avais juste une remarque. Vous disiez un moment que l'utilisation des vignettes semblait favorable au produit, et c'est notamment marqué page 20 dans le tableau. Je me disais qu'on ne pouvait peut-être pas aller aussi loin en disant que c'était favorable aux produits parce que c'est en comparaison de l'analyse de scénarios qui repose sur les données de l'EQ-5D. Mais comme on l'a dit, les données de l'EQ-5D sont peu robustes, il y a un manque d'effectifs, parfois c'est contre-intuitif. Ne pourrait-on pas se limiter à dire, comme vous l'avez dit dans la réserve, que toute cette analyse de la qualité de vie est entachée d'une

incertitude forte, mais sans dire nécessairement que l'analyse des vignettes est favorable au produit ? C'est page 20, chez moi.

Un chef de projet du SEM.- Nous avons projeté le PowerPoint, mais en effet, c'est repris dans l'analyse critique, de la même manière dans le projet d'avis. En effet, c'est vrai qu'on peut peut-être s'en tenir à ce que vous proposez pour ne pas se prononcer sur le fait que ce soit favorable au produit évalué, dans la mesure où on a montré aussi les limites de la construction des questionnaires génériques dans ce cas. On peut peut-être s'en tenir à une incertitude forte sur l'estimation des scores d'utilité dans le contexte spécifique de l'ASMD, et donc sur les résultats en termes de QALY. Ce sera modifié.

M^{me} PARIS.- Merci Fanny. Je voyais une question posée dans le tchat : la CEESP peut-elle faire une remarque sur le design des essais cliniques des maladies rares ? C'est une suggestion qui ne me paraît pas complètement déplacée, vu les reproches que l'on fait aujourd'hui à cette étude. Je ne sais pas si quelqu'un est d'avis que ce n'est pas approprié. Hassan, je ne sais pas si c'est à ce sujet ou non.

M. SERRIER.- Si, c'est à ce sujet. Cela ne me semble pas du tout inapproprié. Par contre, il faut bien faire attention à la formulation si c'est compris dans le design. Ce n'est pas vraiment le design qu'on remet en cause. Ce que j'avais compris, c'est que c'était l'idée qu'en s'y prenant un peu mieux, on pourrait inclure plus de patients. C'est tout, mais ce n'est pas forcément une remarque sur le design des essais qui doivent être menés. Il ne me semble pas. Si c'est compris comme cela, il faut la modifier, mais il ne me semblait pas.

M^{me} PARIS.- Il y a les essais cliniques, et il y a peut-être aussi l'étude en tant que telle. Par exemple le recours à des vignettes, on sait bien qu'en principe, on ne l'accepte pas. Le guide est assez clair là-dessus.

M. SERRIER.- En relisant la phrase, effectivement, il y a une partie où on parle du design des essais cliniques. Il me semblait qu'on avait discuté de l'idée que ça permettait d'inclure plus de patients. Peut-être que je dis une erreur, mais je n'ai pas souvenir sur le design. Que veut-on dire par la réflexion sur le design ?

M^{me} PARIS.- Je laisse parler Driss et Claire parce que c'est sans doute à ce sujet, comme cela, ça nous permettra d'avancer.

M. BERDAÏ.- Peut-être Claire en premier en tant que rapporteur.

M^{me} GRANON.- La discussion était venue du fait qu'on avait une population de personnes traitées actuellement en France, qui est de l'ordre de la trentaine et qu'on n'avait eu que quatre participants français dans l'essai clinique. On se disait tout de même que c'était dommage, que

les essais cliniques ne mobilisent pas plus de personnes, alors qu'il y a une attente importante. C'était un peu le départ de la réflexion. Effectivement, ce n'est pas vraiment le design, on n'a pas dit au départ qu'il ne faut pas qu'il y ait plus de 4 Français, mais c'est un peu le problème du recrutement dans les essais cliniques. C'était la base de notre discussion.

M^{me} PARIS.- D'accord. C'est plutôt le nombre.

M. BERDAÏ.- Pour rebondir sur ce point, effectivement, je serai prudent, et notamment dans la formulation et la compatibilité avec l'avis de la commission de Transparence. Parce que si j'ai bien suivi, c'est tout de même un essai randomisé en double aveugle, comparatif. Le design, d'un point de vue méthodologique, sur les qualificatifs que je viens de dire, il n'y a pas que cela qui compte, mais c'est tout de même un modèle. Dans la formulation, déjà, dans le principe, par rapport à ce que je viens de dire sur l'essai randomisé, contrôlé et double aveugle.

Deuxième point, c'est la compatibilité qu'il faut absolument assurer par rapport à l'avis de la commission de transparence sur les critiques relatives à l'essai clinique.

Le troisième point, c'est sur le recrutement. Tout dépend au tout départ, quel a été le calcul du nombre de sujets nécessaires. Si effectivement on peut reprocher à ce qu'il y ait un recrutement faible en France, il faut faire attention si le nombre de sujets nécessaires calculés en vue de l'essai clinique était un nombre qui a été atteint, quitte à avoir peu de recrutements en France parce que les investigateurs étaient peut-être moins impliqués ou pas assez rapides, ou les procédures de l'essai clinique ont fait que le démarrage était plus tardif qu'ailleurs, mais si le nombre de sujets nécessaires global attendus a été vérifié, là aussi, je serai prudent.

Mais je laisse maintenant commenter par rapport à un dossier que je connais bien moins que bien évidemment les rapporteurs ou les évaluateurs.

M^{me} PARIS.- Je suppose que c'est une notion qu'on n'a pas. Peut-être que c'est dans l'avis de la transparence, mais en général, les détails sur le nombre de patients attendus nécessaires pour la démonstration de ce que l'on veut démontrer, ce n'est pas des données que nous avons habituellement.

M. BERDAÏ.- Exact. Mon point principal est d'être parfaitement compatible avec l'avis de la commission de transparence sur ces aspects.

M^{me} PARIS.- Tout à fait, c'est un très bon réflexe. Christophe, tu voulais dire quelque chose ?

M. DUGUET.- Dans la formulation, cela me gêne de dire qu'il aurait fallu inclure un maximum de patients. Cela veut dire quoi ? Que les 90 patients français auraient dû être mis dans l'essai clinique ? Soit on identifie, comme cela vient d'être dit par Driss, qu'ils n'ont pas mis le bon nombre de patients permettant à l'essai clinique de pouvoir avoir a priori un résultat. Mais dans

ce cas, on peut aussi s'interroger du fait que l'ANSM ait autorisé l'essai. Mais leur dire de mettre le maximum, je ne trouve pas que cela veut dire grand-chose. Le maximum, c'est tous les patients ? On est en plus dans le cadre d'une maladie très rare.

J'ajouterai en plus que s'il n'y a que 4 recrutements en France, je ne connais pas le détail de ce dossier, mais je le connais pour beaucoup de maladies rares, ces essais qui sont compétitifs, malheureusement, dans beaucoup de cas, les lourdeurs administratives des hôpitaux en France font que les centres français arrivent après d'autres centres. C'est aussi pour cela qu'il n'y a pas de recrutement parfois suffisant, pas dans toutes les maladies, mais dans certaines. C'est aussi difficile de le reprocher à l'industriel. Du coup, je trouve que cette phrase devient un peu bancal. Soit on identifie clairement un truc très précis que l'industriel n'a pas fait, par exemple, mais un truc générique de dire qu'il faut mettre le maximum de patients, franchement, cela n'est pas utile.

M^{me} PARIS.- D'accord. Je comprends. C'est vrai qu'on est gêné dans l'évaluation à plusieurs reprises par le fait qu'il n'y ait pas assez de patients, aurait-il dû et pu en recruter plus ? C'est plus compliqué. Un chef de projet voulait parler.

Un chef de projet du SEM.- Merci Valérie. Christophe et Driss ont déjà résumé une partie de mon propos. Je suggère, si nécessaire, de reformuler, éviter le mot design et les aspects méthodologiques, pour les aspects dits par Driss et Christophe d'une part. D'autre part, peut-être penser à une phrase en termes suggestifs sur le recrutement, mais vraiment une phrase qui n'a rien à voir avec la méthodologie.

Pourquoi ? Là, j'aborde un aspect hyper important depuis l'historique de la CEESP et la nature de ses suggestions. C'est que souvent, dans l'évaluation économique, si l'on pense à proposer un design d'un essai, c'est qu'on le propose à partir d'une méthodologie. Cette méthodologie est souvent très sévère et très difficile, non utilisée aujourd'hui. Parce que si l'on propose un design, c'est qu'il faut proposer des instruments de l'analyse de l'incertitude, de la valeur de l'information, du design. On ne l'utilise pas aujourd'hui. C'est la fenêtre de l'évaluation économique par le simple *size*, pour le nombre de sujets, c'est à partir de ces propres instruments intrinsèques à son évaluation.

Je rejoins le raisonnement de Driss, c'est de ne pas rentrer dans ces détails parce que cela fait partie de l'ANSM, de l'EMA sur les designs.

En revanche, s'il y a une phrase qualitative qui peut enrichir la discussion en termes qu'un nombre de sujets..., mais à éviter tout de même, parce que cela soulève toujours les questions de l'évaluation économique. Parce qu'on peut avoir des design dans l'évaluation économique à 3 000 patients et le design n'est pas bon. C'est donc vraiment un piège de parler du nombre de sujets.

On l'a évité à plusieurs reprises, à la CEESP à la fin, parce qu'on a dit : si ça se trouve, il n'y a pas que le design qui pose un problème, et on rentre en conflit avec les avis de la CT, de l'EMA, des associations des patients, et ainsi de suite.

Je rejoins le point de vue de Driss, Christophe et Hassan au niveau de la reformulation.

M^{me} PARIS.- Mon impression à la suite de cette discussion, c'est que le plus simple serait peut-être de supprimer carrément ce paragraphe, même si nous avons été gênés. C'est vrai que l'évaluation économique était entravée par le fait qu'on manque de patients, mais on manque peut-être un peu de solidité pour se prononcer en tant que CEESP.

C'est vrai que l'ANSM ou la commission de transparence, etc., aurait pu faire des remarques sur les essais que peut-être on est un peu moins bien placés pour faire. Voilà mon sentiment.

Je vois que Fanny et Lionel souhaitent se prononcer à ce sujet ou à un autre.

M^{me} MONMOUSSEAU.- Je voulais rebondir sur la remarque de Lionel de tout à l'heure, mais peut-être qu'il voulait réagir lui-même à sa propre remarque, c'était sur le fait qu'il y ait neuf états dans le modèle, finalement. Parce que c'est une maladie rare, il y a un petit effectif et on a une difficulté à alimenter les neuf états du modèle. C'est plus une remarque qui ferait la liaison avec le nombre d'états dans le modèle.

Un chef de projet du SEM.- Sinon, peut-être qu'il faudrait qu'on déplace l'un des arguments qu'on a pu mettre dans l'avis d'efficience comme commentaire de la conclusion de la CEESP. C'est ce qu'on vous projette.

M^{me} PARIS.- C'est à replacer dans la conclusion, c'est ça ?

Un chef de projet du SEM.- Oui, à replacer dans la conclusion en disant : dans le contexte de pathologies rares, la CEESP souligne qu'il faut mener une réflexion sur la structure du modèle la plus adaptée pour pouvoir estimer de façon robuste les probabilités de transition.

M^{me} PARIS.- Lionel, c'était à ce sujet ?

M. PERRIER.- Oui. Simplement, notre réflexion initiale était une remarque générale. C'est-à-dire qu'il faut prêter effectivement attention au moment de la conception et de l'organisation des études et même du suivi, je dirais. Parce que comme le mentionnait à fort juste titre un chef de projet, avec des essais avec plusieurs centaines, voire plusieurs milliers de patients, à partir du moment où il y a des données manquantes, on peut mettre en œuvre de l'imputation multiple et essayer de trouver des résultats les moins biaisés possible. Sur un faible échantillon, on ne peut pas mettre en œuvre de l'imputation multiple. Ce sont des contraintes supplémentaires. Mais là

encore, même si je partage bien entendu les réflexions qui viennent d'être faites, c'était vraiment une réflexion générale sur les essais.

M^{me} PARIS.- La phrase que propose en direct les chefs de projet est-elle suffisante ? Comprend-on bien qu'on fait référence aux neuf états ? Ou préfère-t-on ne pas faire référence aux neuf états ?

M. PERRIER.- Il me semble qu'il faut tout de même éviter que ce soit contre-productif. C'est-à-dire qu'on sait bien que les avis ne sont pas lus que par l'industriel, mais aussi par tous les autres. Là, un industriel qui a fait un effort d'essayer de décrire le mieux possible la maladie, si l'on envoie le message en disant : non, il faudrait faire beaucoup plus simple, on risque après, pour d'autres dossiers d'autres industriels, de dire votre modèle est beaucoup trop simple et il ne reflète pas l'évolution de la maladie. Attention tout de même.

M^{me} PARIS.- J'entends bien. En même temps, neuf états, avec très peu de patients, cela laisse très peu de chances d'arriver à construire quelque chose de robuste.

Premièrement, laisse-t-on une remarque assez générale comme celle qui figure là à l'instant à l'écran ? Je lis : dans le contexte spécifique de pathologies rares, avec peu d'observations dans les essais, la CEESP souligne qu'une réflexion sur le type de modèle développé doit être menée afin de pallier à certaines limites liées à l'estimation des probabilités de transition et des scores d'utilité. Je ne sais pas si ce que l'on veut dire est très clair ?

Un chef de projet du SEM.- C'est vrai que l'intention est bonne et j'adhère sur le fond de la chose, mais c'est très engageant. Je rejoins un peu l'avis de Christophe, parce qu'on a peu d'observations, le type de modèle. Cela fait dix ans qu'on connaît la problématique statistique depuis la création de l'évaluation économique, après les premiers travaux de 1983, les travaux pionniers de Riemontais et Briggs. Quel est le type de modèle ? Je ne sais pas. C'est un peu engageant. On a quatre ou cinq types de modèles. Cela fait trop technique et ça peut donner une information aux industriels : on va engager un type de modèle.

Or c'est à éviter ou peut-être soulever l'importance... Je suis un peu sceptique là-dessus, à l'instar de tout ce qui s'est fait au NICE, de ne pas trop dire là-dessus. Parce que les maladies rares, très rares, nous n'avons pas vraiment de consensus, notamment méthodologiques. Parfois, il y a même des remises en question du cadre classique. C'est un terrain un peu mitigé méthodologiquement.

M^{me} PARIS.- Driss.

M. BERDAÏ.- Je voulais juste dire, même si ce n'est pas mon domaine d'expertise, que soit on est suffisamment claire pour qu'effectivement l'industriel comprenne sans ambiguïté ce qui est exprimé par la commission, c'est-à-dire, si j'ai bien compris, l'inadaptation du niveau de détail du modèle par rapport aux données disponibles. Soit, on supprime ce paragraphe. Il n'a de valeur qu'à partir du moment où il est bien compris par le destinataire.

Un chef de projet du SEM.-Exactement.

M. BERDAÏ.- Tout en appréciant l'effort fait dans la construction du modèle, mais en constatant peut-être qu'en la circonstance et par rapport au nombre de données disponibles, il est trop complexe.

M^{me} PARIS.- Je voudrais juste une précision, parce que j'ai l'impression que vous enlevez le paragraphe de là où il était pour le mettre dans la conclusion. On peut tout de même réfléchir en deux étapes, c'est-à-dire : a-t-on envie de dire quelque chose là-dessus ? Cela ne me paraît pas complètement aberrant, malgré tout, de dire quelque chose là-dessus. Parce qu'on s'est tout de suite dit les neufs états, c'est compliqué à renseigner.

Il y a aussi certaines données sur l'utilité, évidemment on ne les utilise pas, on ne les garde pas, mais qui laissent penser que finalement, il y a plusieurs états qui, en termes de qualité de vie, etc., ne sont pas forcément très discriminés. Je trouve que se poser la question n'est pas aberrant maintenant, donc je la laisserai là où elle était.

Maintenant, veut-on la mettre en conclusion, ce qui lui donne une portée un peu différente ? C'est un peu la question que l'on se pose tout de suite. Peut-être, Lionel, puisque tu suggérais de le mettre en conclusion, après toutes les discussions qu'on a eues, qu'en déduis-tu ?

M. PERRIER.- Personnellement, je resterai sur l'option initiale.

M^{me} PARIS.- C'est-à-dire qu'elle n'est pas dans la conclusion. On laisse la phrase : il n'est pas à exclure que le développement d'un modèle... les plus pertinents cliniquement, aurait pu permettre de pallier certaines limites est lié à l'estimation des probabilités de transition et des scores d'utilité. Cette phrase te convient-elle ? Mais elle est dans la conclusion ?

Un chef de projet du SEM.- On va l'enlever.

Un chef de projet du SEM.- Non, elle n'est pas dans la conclusion, elle est dans : en ce qui concerne la conformité méthodologique.

M^{me} PARIS.- D'accord.

Un chef de projet du SEM.- C'était sa position initiale.

M^{me} PARIS.- D'accord. C'était position initiale. Lionel ?

M. PERRIER.- Je laisserai tel quel, c'est-à-dire dans le corps du texte, mais pas forcément en conclusion.

M^{me} PARIS.- D'accord. On enlève en conclusion. Hassan.

M. SERRIER.- J'allais aller dans le même sens, surtout qu'il y a un point important, c'est qu'on est moins affirmatif dans la formulation telle qu'elle est dans le corps actuellement, que celle qu'on avait mise la conclusion. On lance une piste : cela aurait pu être intéressant de se poser la question si un modèle plus simple aurait aidé, et sans le mettre en avant dans la conclusion. Cela me convient très bien comme ça.

M^{me} PARIS.- D'accord. Peut-on s'entendre sur cette option ? Y a-t-il des objections ? Y a-t-il d'autres commentaires ? Très bien. Je propose que l'on passe au vote.

(Il est procédé au vote par appel nominatif.)

Résultat du vote :

Favorable : 20 voix

Défavorable : 0 voix

Abstention : 0 voix

(L'avis est adopté à l'unanimité.)

M^{me} PARIS.- Merci beaucoup à tous pour le traitement de ce dossier.