

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.

- **Validation d'un avis économique - Avis économique - LIBMELDY indiqué dans le traitement de la leucodystrophie métachromatique (LDM) caractérisée par des mutations bialléliques du gène de l'arylsulfatase A (ARSA) entraînant une réduction de l'activité enzymatique de l'ARSA : chez les enfants atteints de la forme infantile tardive ou juvénile précoce, sans manifestations cliniques de la maladie ; chez les enfants atteints de la forme juvénile précoce, présentant des manifestations cliniques précoces de la maladie, qui ont conservé la capacité de marcher indépendamment et avant l'apparition du déclin cognitif (ORCHARD THERAPEUTICS) - ECO-EFFI 774**

M^{me} CHEVREUL.- On redémarre.

Une cheffe de projet, pour la HAS.- On a beaucoup de choses à vous dire aujourd'hui, on a des interrogations, et on a besoin de discussions collectives. C'est le dossier LIBMELDY qu'on appellera ati-cel.

Au niveau du contexte, ati-cel a obtenu une AMM en décembre 2020 dans la leucodystrophie métachromatique de l'enfant, dans les formes infantiles tardive et juvénile précoce asymptomatiques et également la forme juvénile précoce présentant des manifestations cliniques précoces de la maladie qu'on appelle également paucisymptomatiques.

En 2021, la commission de transparence avait évalué le produit et accordé le remboursement uniquement dans la population asymptomatique. La commission avait demandé de disposer dans un délai maximal de trois ans des données de suivi des études en cours. Faute d'accord sur le prix dont le coût unitaire revendiqué, je le rappelle, est d'environ 2 935 000 euros, le produit s'est vu renouveler trois fois son accès précoce et le CEPS a souhaité disposer d'une évaluation économique expertisée par la HAS afin d'éclairer les négociations de prix en cours.

Vous évaluez actuellement la première évaluation économique du produit qui s'inscrit dans le cadre de la soumission de nouvelles données auprès de la commission de transparence. L'industriel a revendiqué un SMR important dans toutes les populations, une ASMR de niveau II dans les populations asymptomatiques et une ASMR de niveau III dans les populations paucisymptomatiques. Le chiffre d'affaires revendiqué est de 10,3 millions d'euros avec une population éligible de 3 à 4 patients.

Dans le cadre de ce dossier, nous avons les contributions de deux associations de patients qui mettent en avant la gravité de la pathologie sur les patients en termes de morbidité, mortalité et de leurs aidants et plaident pour l'accès précoce de la thérapie génique qui présente des effets prometteurs.

La leucodystrophie métachromatique est une maladie génétique ultra rare. On recense environ dix naissances par an en France. Elle se caractérise par une anomalie génétique de l'enzyme arylsulfatase, ce qui provoque l'accumulation de sulfatides dans le système nerveux, central et périphérique, entraînant la destruction de la gaine de myéline autour des nerfs.

Les premiers symptômes peuvent survenir de la petite enfance à l'âge adulte, mais nous allons nous concentrer uniquement sur les formes juvéniles, les formes dites précoces pour lesquelles ati-cel est recommandée. Il y a la forme infantile tardive qui commence avant l'âge de 30 mois et la forme juvénile précoce, dont les premiers symptômes surviennent entre 30 mois et 6 ans.

Sur ce graphique, vous pouvez voir l'évolution de la maladie. En ordonnée, on y reviendra, se trouve la classification GMFC, qui est une classification des fonctions motrices qui va de 0 : la normale, 2 : la perte de la marche autonome, 6 : correspondant à une perte de la tenue de la tête.

L'évolution entre les deux formes est assez similaire avec un développement initial normal. On observe un ralentissement, puis une stagnation des acquisitions avec une dégradation extrêmement rapide et sévère des fonctions motrices cognitives et sensorielles qui conduisent en quelques semaines à quelques mois à un état grabataire et au décès précoce des enfants.

Dans la forme juvénile précoce, la particularité est que la phase de plateau est un peu plus longue, mais dès que la marche est perdue, le déclin est aussi rapide que dans l'autre forme. La survie est également un peu plus longue, mais les patients se trouvent dans une situation de handicap extrêmement lourde.

Par rapport au mécanisme d'action d'ati-cel, il s'agit d'un médicament de thérapie génique qui repose sur l'autogreffe de cellules souches hématopoïétiques génétiquement modifiées pour produire l'enzyme ARSA fonctionnelle et rétablir la production de l'enzyme ARSA à des taux normaux ou supra-physiologiques qui seraient durables au cours du temps. L'image en bas de la diapositive montre cette surexpression de l'enzyme ARSA qui est rétablie avec une stabilisation de la production dès trois mois et une production durable au cours du temps sur le suivi disponible.

Cette diapositive est essentiellement pour préciser le fait qu'ati-cel est spécifique à une population ciblée de par les symptômes et les critères d'éligibilité et également s'administre à l'aide d'un circuit très spécifique et très spécialisé dans des structures expertes. Il y a effectivement des enjeux que l'on a identifiés sur l'identification des patients éligibles et le profil de ceux traités en pratique, ne serait-ce que par la possibilité parfois d'avoir une évolution de l'état de santé du patient entre le moment où il est éligible et le moment où il reçoit finalement le traitement.

Au niveau du programme de développement, il a été réalisé dans un seul centre à Milan depuis près de dix ans. Les patients venaient du monde entier dont un patient français. Tous les patients ont été évalués de la même manière qu'ils soient traités ou non par ati-cel, ce qui est un point important de ce dossier. Le programme s'appuie sur une série de données de patients non traités issus d'une cohorte historique de 43 patients, et un total de 39 patients traités par ati-cel dans deux études cliniques de phase I, II non randomisées en ouvert et trois programmes italiens d'accès étendu.

Pour rappel, trois informations importantes concernent les similarités entre les différentes études du programme de développement. La première concerne le co-critère de jugement principal, le différentiel d'effet versus la cohorte historique a été uniquement évalué sur le score GMFM, qui était le critère principal, avec un appariement qui a été mené selon l'âge des patients. Le deuxième élément, c'est de vous signifier que l'autre essai de phase I, II reposait sur la forme cryoconservée, alors que l'essai pivot reposait sur la forme fraîche, et c'est la forme cryoconservée qui est celle qui est commercialisée en circulation actuellement. Le troisième élément à souligner, c'est le suivi médian. Effectivement, il y a un suivi assez substantiel pour ce type de pathologie, avec un suivi maximal d'environ 11 ans, mais il faut rappeler que parmi les patients traités, un peu moins de la moitié a un suivi inférieur à cinq ans.

Les deux outils d'évaluation des fonctions motrices utilisés dans les essais cliniques sont le GMFM et le GMFC. Le critère principal, le GMFM est un score composite sur cinq dimensions. Le score va de 0 à 100 %, et c'est l'un des critères principaux de l'essai pivot sur lequel l'effet d'ati-cel a été évalué par rapport à la cohorte historique. Nous avons en parallèle le GMFC, qui est une classification, utilisée comme critère secondaire dans la modélisation économique par l'industriel pour modéliser l'effet d'ati-cel. L'essai pivot, dans son protocole, ne prévoyait pas de démonstration d'un effet d'ati-cel versus la cohorte historique sur ce critère secondaire.

La méthode d'appariement sur l'âge a été menée sur le critère principal. Il s'agissait d'un appariement non exact uniquement sur l'âge qui consistait à inclure les patients non traités de la cohorte historique uniquement si leur âge était compris entre une fourchette définie par l'âge le plus bas, moins trois mois, et l'âge le plus élevé des patients traités. Si l'âge des patients non traités était inclus dans cette borne, ils étaient inclus dans la population appariée, et c'était sur cette population qu'était estimée l'efficacité relative. Comme je le rappelais, cet appariement n'a pas été mené sur le score de classification utilisé dans la modélisation, le GMFC.

Nous allons vous présenter les résultats de l'essai pivot, tout en précisant qu'avec les rapporteurs, nous avons eu quelques difficultés à identifier les résultats cliniques des études, puisqu'ils n'ont pas été présentés dans le dossier soumis à la CEESP. Sur le critère principal, l'industriel a

uniquement fait un renvoi au CSR et au dossier soumis côté CT, ce qui est fortement regrettable. Nous rappelons également que ces données en sous-population sont de nature exploratoire, car il n'y avait pas de gestion du risque alpha, néanmoins, nous voulions les présenter, car il nous a semblé important d'avoir comme base de discussion les résultats sur le critère principal de l'essai pivot.

Ceci étant dit, sur les huit patients asymptomatiques, nous observons une amélioration de la motricité et du score GMFM à deux ans et trois ans chez les patients traités, en bleu, par rapport à l'évolution naturelle de la maladie, en vert. Sur le moyen terme, on voit que la majorité des patients maintiennent des performances motrices avec malgré tout des profils de réponse variables d'un patient à un autre. Certaines courbes individuelles montrent une baisse progressive, notamment la dernière courbe qui chute.

Sur le graphique en bas à gauche, qui représente la survie sans perte de locomotion indépendante, tous les patients, sauf un, avaient maintenu la capacité de marcher en autonomie, alors que les patients non traités avaient une perte très rapide des fonctions motrices ou étaient décédés.

La particularité de ce plan de développement était que les patients inclus et traités dans l'étude pivot pouvaient avoir des frères et sœurs atteints de la pathologie, mais non traités, qui étaient inclus dans la cohorte historique. La comparaison des frères et sœurs traités et non traités est particulièrement intéressante, puisqu'ils vivent dans le même environnement. On pouvait voir que, quand les fratries existaient, les enfants traités avaient des fonctions préservées par rapport à leurs frères et sœurs non traités.

Si l'on regarde les résultats sur la forme juvénile précoce asymptomatique avec un effectif de quatre patients, la diapositive reprend les mêmes données. Sur le premier graphique, on voit une préservation des fonctions motrices à deux ans et trois ans ; sur le moyen terme également, des signaux positifs avec une réponse néanmoins variable pour un patient et également, en bas à gauche, une préservation de la locomotion indépendante chez tous les patients, sauf un. De la même manière, les bénéfices étaient aussi confirmés lorsqu'on regardait les frères et sœurs traités et non traités.

Pour les résultats sur la sous-population juvénile précoce paucisymptomatique, on observe un différentiel d'effets, mais il y a une incertitude plus importante qui est portée sur cette sous-population au niveau des résultats et également lorsqu'on regarde la courbe en bas à gauche, où cette fois-ci, les courbes individuelles étaient parfois très superposables avec la cohorte historique.

Voilà pour les résultats cliniques, en sachant qu'il y a aussi le chef de projet médico-technique présent avec nous, qui pourra éventuellement répondre à des interrogations. Nous allons passer à l'analyse de l'efficience.

Un chef de projet, pour la HAS.- Merci. L'industriel n'a présenté qu'une analyse coût-efficacité exprimée en euros par année de vie gagnée, car il ne disposait pas de données de qualité de vie pertinentes et robustes, que ce soit dans l'essai clinique ou dans la littérature, qui auraient permis de fournir une analyse coût-utilité. La perspective retenue est celle du système de santé, l'horizon temporel est quant à lui limité à 30 ans.

Nous avons trois sous-populations d'analyse représentant chaque sous-type concerné par l'indication. Le comparateur est constitué des meilleurs soins de support, car il n'existe malheureusement aucun traitement. La population simulée est la population MITT. Quelques patients ont été exclus, car ils n'étaient pas éligibles à l'indication retenue par l'EMA.

Le modèle retenu est un modèle de Markov à huit états de santé définis par des scores GMFC. Cela correspond à une échelle de la fonction motrice d'un niveau 0 à 6. Le patient ne peut transiter que vers des états GMFC plus sévères et l'état GMFC 6 précède l'état de décès. Le score GMFC représentait un critère secondaire exploratoire dans l'essai clinique pivot, 201222.

Le choix de ce critère est argumenté par le fait qu'on peut y associer des états de santé et une prise en charge spécifique pour chaque niveau. L'industriel a argumenté qu'il ne pouvait pas faire avec le score GMFM.

Toutefois, ce critère peut poser problème parce qu'aucun effet statistiquement significatif n'a été démontré sur ce critère GMFC. Une réserve mineure est d'ailleurs proposée en lien avec cette limite, notamment en raison de l'absence d'utilisation du critère principal au sein d'une autre structure de modèle qui n'est pas suffisamment justifiée à notre sens.

La modélisation de l'efficacité se fait en deux étapes. D'abord, il faut comprendre que le parcours des patients dans le modèle dépend de leur type de réponse. Les patients sont classés selon leur type de réponse à MLDI, ce qui va constituer la première étape. Le type de réponse est défini comme "répondeur complet" s'ils sont stables sur le plan moteur et cognitif et restent dans l'état GMFC 0. Cela implique qu'ils soient traités avant l'apparition des symptômes, ce qui exclut donc tous les patients juvéniles paucisymptomatiques. Tous les patients paucisymptomatiques ne peuvent pas être répondeurs complets. Les patients sont classés comme "répondeurs partiels stabilisés" soit s'ils sont traités après l'apparition des symptômes et qu'ils se stabilisent par la suite, ou qu'ils montrent une progression après l'administration de MLDI, mais qu'ils se stabilisent ensuite. Enfin, les "répondeurs partiels non stabilisés" sont les patients qui sont traités, mais qui présentent une aggravation continue des fonctions motrices et cognitives. Cette progression est

tout de même plus lente que celle des patients non traités. Ce sera modélisé dans le modèle de Markov par l'application d'un multiplicateur d'effets, mais on va détailler cette partie un peu plus tard.

Ensuite, la deuxième étape consiste en l'application de multiplicateurs d'effets uniquement chez les patients qui ne sont pas stabilisés et qui sont fondés sur le critère secondaire GMFC dont on vient de discuter à la diapositive précédente.

La classification des patients selon le type de réponse est fondée sur plusieurs critères cliniques : le quotient intellectuel, une IRM, le score GMFM et le score GMFC. Notre analyse clinique révèle la fragilité de ces données, car la classification repose sur une évaluation ad hoc de l'essai clinique et repose sur des données patients avec des durées variables, ce qui entraîne un risque de biais. De plus, la classification repose sur des hypothèses cliniques, comme le fait qu'un patient va être considéré comme stable malgré une progression du score GMFC si le score GMFM et d'autres marqueurs restent stables.

Il est aussi important de souligner que cette classification constitue le paramètre clé de la modélisation, car c'est ce qui va influencer l'évolution clinique des patients. Vous le verrez lors de la présentation des résultats, mais cela va se traduire par un impact important de ce paramètre observé à travers les analyses de sensibilité déterministes. Au regard de tous ces éléments, nous vous proposons une réserve importante sur ce paramètre.

Toujours concernant ce paramètre, nous avons des hypothèses qui sont sous-jacentes relatives au maintien de l'effet traitement sur l'horizon temporel. La première hypothèse est que la durabilité de l'effet du traitement pendant huit ans est soutenue par un rationnel clinique relatif au mécanisme d'action d'ati-cel, et une durée médiane de suivi de 7,46 ans dans la population juvénile précoce et de 9,52 ans dans la population infantile tardive.

Cette hypothèse présente des limites, car elle repose sur des données limitées et une incertitude demeure sur la progression ou la stabilisation des patients à plus long terme. D'ailleurs, certaines études externes à l'essai pivot qui ont été mobilisées ont une durée médiane de suivi inférieure à trois ans.

Après huit ans, une hypothèse de réduction de l'effet du traitement est appliquée chez 0,02 % des patients à chaque cycle, c'est-à-dire tous les mois. Ces patients se verront appliquer des probabilités de transition des patients non stabilisés. Mais cette hypothèse repose sur un avis d'expert pour une autre thérapie génique. Même si l'applicabilité de ce taux est incertaine, l'approche reste plus conservatrice que de ne pas prendre en compte du tout de perte d'effet.

Les patients qui ne sont pas stabilisés se voient appliquer une probabilité de transition fondée sur la durée de séjour passée dans chaque niveau de score GMFC. L'estimation est réalisée à partir de données de la cohorte historique NHX et des données d'ati-cel issues de l'essai pivot et des autres études externes. Pour calculer les probabilités de transition du bras traité par ati-cel, l'industriel a utilisé des multiplicateurs d'effets. Ce multiplicateur d'effets n'est rien d'autre qu'un rapport de temps passé dans un état entre les patients traités par ati-cel et les patients de la cohorte historique.

Les trois points sur lesquels nous souhaitons attirer votre attention sont les suivants. L'estimation de ces multiplicateurs d'effets est fondée sur des données agrégées de plusieurs études. Pour l'analyse économique, l'industriel a mobilisé une population plus large que celle retenue dans la démonstration de l'efficacité de l'essai pivot. Deuxièmement, dans le modèle d'efficacité, l'efficacité est estimée à partir d'un critère secondaire exploratoire, le score GMFC. Troisièmement, contrairement à l'essai pivot, aucune méthode d'ajustement n'a été mise en place pour tenir compte de la variabilité inter-individuelle.

Plusieurs hypothèses ont été formulées par l'industriel. C'est sur ces hypothèses que l'on souhaite attirer votre attention et qu'on souhaitera ensuite avoir une discussion avec vous.

La première porte sur l'extrapolation de l'effet observé sur le critère principal GMFM au critère secondaire GMFC. Malgré une absence de démonstration d'efficacité dans l'essai pivot sur le critère de jugement secondaire fondé sur le score GMFC, l'industriel fait l'hypothèse que l'efficacité observée du critère principal le score GMFM, dans l'essai pivot, peut être extrapolée au score GMFC. Cette hypothèse est principalement argumentée par l'industriel par un test de Spearman qui révèle une corrélation inverse entre les deux scores. La corrélation de Spearman n'est pas une mesure de concordance et n'implique pas une corrélation entre les effets du traitement sur ces deux critères.

Nous vous proposons donc une réserve importante, car la relation de causalité entre l'effet observé sur le critère principal et le critère secondaire utilisé dans la modélisation n'a pas été démontrée.

Je vais laisser la parole à notre cheffe de projet pour vous introduire le contexte de la deuxième hypothèse qui serait potentiellement sujette à une réserve majeure.

Une cheffe de projet, pour la HAS. - Merci. La corrélation évoquée par l'industriel n'implique pas nécessairement une causalité d'effet entre les deux critères.

Le deuxième niveau où l'effet d'ati-cel est intégré concerne l'estimation des probabilités de transition des patients non-stabilisés. Ces patients non-stabilisés traités par ati-cel se voient

appliquer une progression moins rapide dans les différents états de niveau GMFC par rapport à la cohorte historique. Un effet est bien appliqué à ces patients. Cet effet est calculé à partir de multiplicateurs, sorte de proxy d'un effet relatif. Ces multiplicateurs sont calculés en comparant les données brutes du temps passé dans chaque niveau de GMFC entre ati-cel et le *best supportive care*, sans méthode assurant la comparabilité des populations entre les groupes évalués, sachant que nous sommes dans ce contexte d'essais non-randomisés. Il existe probablement un risque de biais dans l'estimation de l'effet d'ati-cel, ce qui pourrait conduire à une estimation biaisée du RDCR.

L'objectif aujourd'hui est de vous présenter les arguments de l'industriel et les éléments de réflexion qu'il apporte, ainsi que les hypothèses justifiant son approche, de proposer ensuite notre analyse critique et dans un second temps d'essayer d'en tirer une conclusion collective sur le niveau d'acceptabilité des hypothèses formulées et le niveau de réserve la plus adéquate.

L'un des arguments importants de l'industriel est de dire que l'appariement sur l'âge, tel qu'il a été effectué sur le critère principal, n'est pas adapté pour l'évaluation économique basée sur le critère secondaire. L'industriel explique que l'appariement sur l'âge à un point temporel de deux ou trois ans fournit une distribution de la proportion de patients dans chaque classification à un temps précis, mais pas des durées de séjour dans les différents états. Ces durées de séjour sont nécessaires pour alimenter le modèle économique. De plus, la méthode d'appariement sur l'âge réduirait la population appareillée, rendant difficile ensuite l'estimation robuste des durées de séjour moyennes.

Ces deux premiers arguments concernent l'impossibilité de mise en œuvre d'un point de vue technique. Le troisième argument, plus clinique, soutient que l'appariement sur les frères et sœurs était le seul méthode logiquement faisable et pertinent dans le cadre de cette pathologie.

Vis-à-vis de ces premiers éléments, de notre côté, nous estimons que ces justifications sont partielles et entendables. Néanmoins, cela soulève des questions plus fondamentales sur le choix du critère retenu pour modéliser l'effet d'ati-cel dans la modélisation et sur le choix de la méthode utilisée pour estimer un effet d'ati-cel non biaisé. Si l'appariement sur l'âge n'était pas possible, n'y aurait-il pas d'autres méthodes pour limiter les biais dans la comparabilité des populations, applicables au critère retenu dans la modélisation ? Sinon, si ce n'est pas possible, pourquoi ne pas avoir utilisé le critère principal sur lequel, dans l'essai pivot, un appariement sur l'âge semble avoir été fait et soumis aux autorités ?

L'industriel a peu documenté l'impossibilité de réaliser l'appariement sur l'âge sur le critère secondaire, dans le sens où il est très compliqué de savoir si cela était réellement impossible. "Non pertinent" ne signifie pas nécessairement "non possible". De plus, l'argument selon lequel

l'appariement réduirait l'effectif n'est pas une raison valable. Ce n'est pas parce qu'on anticipe une limite qui pourrait survenir d'une méthode adéquate mise en place que c'est une raison valable pour ne pas réaliser cette méthode. A priori, ce serait à la CEESP d'évaluer l'incertitude associée à cette méthode si elle avait été menée et argumentée.

Par ailleurs, si on passe outre cette méthode d'appariement sur l'âge, l'industriel n'a pas proposé d'études de faisabilité pour d'autres méthodes d'ajustement statistique, comme la pondération ou l'appariement. Résultat, il est très compliqué de savoir si réellement il n'y avait pas une méthode plus robuste qui aurait pu être proposée et menée.

À ce stade de l'analyse, le fait de ne pas avoir mené une méthode conforme aux recommandations en vigueur est susceptible d'entraîner une réserve majeure, car la méthode pourrait conduire à une estimation biaisée de l'effet d'ati-cel, et que l'approche réalisée n'est pas convenablement justifiée.

Ensuite, l'industriel rappelle l'efficacité relative démontrée sur le critère principal, avec un niveau de preuve reconnu par les agences sanitaires. Effectivement, ati-cel est une innovation qui peut être qualifiée de rupture, qui a une efficacité assez importante dans certaines sous-populations par rapport à des patients non traités qui décèdent de la pathologie.

L'industriel rappelle ces résultats et souligne que le protocole avait été réalisé sur un seul centre, monocentrique, ce qui pourrait limiter des biais. Il rappelle également que les résultats avaient été appariés sur deux types de contrôles, à la fois sur l'âge et la forme de la maladie, et sur des données appariées frères et sœurs. Il évoque, à partir des données de la littérature, que l'évolution de la maladie est similaire au sein d'un même phénotype. De ces constats, il émet l'hypothèse que l'efficacité relative démontrée dans l'essai pivotale apparié pourrait être transposable aux données simulées dans la modélisation économique. C'est l'hypothèse 2 qui ressort.

Le premier argument qui nous dérange, si je peux le dire ainsi, c'est d'une part que, comme le chef de projet a pu vous le dire, nous nous interrogeons sur la démonstration d'une causalité d'effets entre le critère principal et le critère secondaire. D'autre part, nous nous interrogeons également sur la comparabilité des groupes de patients traités et non traités.

En l'occurrence, voici le tableau des caractéristiques des patients traités et non traités. Ce qu'on peut dire, c'est que la comparabilité n'est pas assurée, notamment en ce qui concerne l'âge au moment des premiers symptômes. On voit que les patients de la cohorte historique avaient des premiers symptômes plus tôt que les patients traités par ati-cel, ce qui pourrait être le signe d'une progression de la maladie plus agressive.

La deuxième problématique est aussi liée à la nature de la cohorte historique. Nous avons très peu de données rétrospectives sur lesquelles on peut comparer les caractéristiques des patients, notamment ne serait-ce que par exemple le quotient intellectuel, le taux d'ARSA ou le niveau GMFC. Par ailleurs, un autre argument qui fait défaut dans ce dossier, c'est que les caractéristiques de la population des 35 patients simulés ne sont pas présentées. On a uniquement les caractéristiques de la population totale traitée par ati-cel.

Vous l'aurez compris, la méthode d'estimation de l'effet d'ati-cel au travers des probabilités de transition sans méthode visant à rendre comparables les populations traitées et non traitées pourrait conduire à l'appréciation d'une réserve majeure au vu des recommandations actuelles.

Néanmoins, on pourrait aller vers une autre direction si on accepte le fait que l'efficacité qui a été démontrée dans l'essai pivotal versus une cohorte historique serait transposable à la population simulée dans le cadre de la modélisation, que ce soit en termes de critères d'efficacité retenus, mais également en termes de population incluse et comparée.

Cela laisse en suspens un certain nombre d'hypothèses que nous souhaiterions discuter collectivement, notamment le fait qu'a priori, si les patients ont été suivis, qu'ils soient traités ou non traités dans un même centre et évalués avec des outils similaires, peut-être que cela réduirait le risque de biais de confusion entre les populations. Il y a une deuxième hypothèse : le fait que dans les essais cliniques évaluant ati-cel, il y avait la forme fraîche et la forme cryoconservée. Il faudrait aussi accepter l'hypothèse que la forme d'ati-cel n'a pas d'impact sur son efficacité. En l'absence d'essais comparant ces deux formes, cela reste une hypothèse.

En termes de biais de sélection, il faudrait accepter que, si l'on se concentre sur l'âge d'apparition des premiers symptômes par exemple, cette différence d'âge entre les groupes comparés ne soit pas de nature à influencer l'effet du traitement au sein d'un même phénotype. Par ailleurs, pour tous les autres facteurs de phénotype ou modificateurs d'effets qui sont inconnus, ceux-ci ne seraient pas susceptibles de compromettre la comparabilité des groupes au sein d'un même phénotype faisant l'occurrence au sous-type de la maladie.

Cela fait beaucoup d'informations. Je ne sais pas si l'on discute tout de suite de la réserve sur l'effet relatif ou si l'on présente les résultats.

Mme CHEVREUL.- Comme vous voulez.

Un chef de projet, pour la HAS.- Pour avoir tous les éléments, il est important de vous présenter les résultats avant de discuter les autres aspects de la modélisation : l'estimation de la mortalité, l'intégration des événements indésirables, les coûts, etc. Ces résultats apporteront des éléments

supplémentaires pour aider à arbitrer le niveau de la réserve de l'hypothèse de la transposabilité de l'effet observé à la population pivotale.

Le RDCR de la population infantile asymptomatique est de 193 635 euros par année de vie gagnée. Les analyses de sensibilité déterministes révèlent que le paramètre ayant le plus d'impact sur les résultats est la classification du type de répondeur, comme je vous le disais tout à l'heure. Par ailleurs, l'analyse de sensibilité en scénario avec un appariement sur les frères et sœurs montre un impact de -5 %.

Le RDCR de la population juvénile précoce asymptomatique est de 242 660 euros par année de vie gagnée. Les analyses de sensibilité déterministes montrent encore une fois que le paramètre qui a le plus d'impact sur les résultats est la classification du type de réponse. Cette fois, l'analyse de sensibilité en scénario avec un appariement sur les frères et sœurs montre un impact de +1 %.

Enfin, concernant la population juvénile précoce paucisymptomatique, le RDCR est de 593 472 euros par année de vie gagnée. Cette fois, les analyses de sensibilité déterministes montrent qu'en plus de la classification du type de réponse, les multiplicateurs d'effets ont aussi un impact sur les résultats. Ceci est logique au regard du fait que la proportion de patients à réponse partielle non stabilisée représente 80 % de cette sous-population contre 8 à 10 % dans les deux autres sous-populations.

On observe aussi que l'analyse de sensibilité en scénario avec un appariement sur les frères et sœurs entraîne une plus forte variabilité avec une baisse du RDCR de 31 %. **Une cheffe de projet, pour la HAS.** - Pour conclure, le tableau récapitulant l'ensemble des réserves proposées concernant l'analyse de l'efficacité vous est présenté sur cette diapositive. Les éléments sont détaillés dans l'avis qui vous a été transmis. On veut bien avoir vos impressions sur toutes les questions posées notamment sur la méthode d'estimation des probabilités de transition.

M^{me} CHEVREUL. - Morgane.

M^{me} MICHEL. - Merci beaucoup pour la présentation. Je voudrais être sûre d'avoir compris quelque chose dans votre argumentaire. Si je comprends bien, vous nous dites qu'il y a un critère de jugement principal qui est la motricité globale, qui a un effet significatif, qui est positif dans l'essai, etc. Il y a un critère secondaire qui évalue aussi, mais pas de la même façon avec le même questionnaire, la motricité globale. C'est un critère secondaire, ils n'ont pas été regarder si c'était une évolution significative. Pour autant, on voit une évolution. Et là, vous nous dites qu'on ne peut pas faire le lien de causalité entre ça et le médicament ? D'un point de vue purement statistique, je suis tout à fait d'accord avec vous, mais avec un Spearman à 0,97, c'est bizarre de dire cette évolution n'est pas en lien. Ou j'ai raté quelque chose.

M^{me} CHEVREUL.- Dans le *process*, voulez-vous répondre ? Avez-vous une réponse ?

Une cheffe de projet, pour la HAS.- Effectivement, d'un point de vue clinique et statistique, on considère que la corrélation est justifiée. Maintenant, on n'est pas certain de la transposabilité du niveau d'effet démontré sur le critère principal et celui appliqué sur le critère secondaire.

M^{me} CHEVREUL.- Par contre, si j'ai bien compris aussi, ils ne pouvaient pas modéliser sur le critère principal parce qu'on continue. Alors que là, ils ont utilisé celui-ci parce qu'il leur permettait d'avoir des états dans le modèle et que c'était beaucoup plus facile d'obtenir des résultats.

Une cheffe de projet, pour la HAS.- Oui, c'est le sujet central : le choix de la structure du modèle et le choix du critère retenu pour modéliser l'effet du traitement.

M^{me} CHEVREUL.- Avant que nous lancions la discussion, je vais proposer aux rapporteurs de nous communiquer leur rapport. Ainsi, nous pourrions en discuter tous ensemble, ce qui éclairera notre échange. Parmi les rapporteurs, il y a Jennifer et Clémence. Vous êtes-vous arrangées sur un ordre ou avez-vous une préférence ? Jennifer, nous vous écoutons.

M^{me} MARGIER.- Je vais être brève. Juste sur ce point du critère, effectivement, on en a pas mal discuté. En tant qu'économiste, ça ne pose pas de problème de discrétiser le critère principal parce que c'est plus facile. Ça ne veut pas dire qu'il n'y avait pas d'autres alternatives possibles, mais effectivement, en discrétisant, on perd de l'information.

Notamment, une des questions qu'on leur avait posées quand ils sont venus en audition, c'est quel est le pourcentage de mal classés en utilisant... On n'a jamais eu de réponse parce qu'ils n'ont peut-être pas compris la question. Je ne sais pas, mais il y a tout de même un risque de mal classés. Il n'est pas négligeable parce que quand vous voyez les intervalles de confiance, ça se chevauche. Ça veut dire que forcément, il y a des mal classés. On perd de l'info et on ne sait pas exactement quel est l'impact.

Mais sur le plan purement économique, ça ne me dérange pas tant que ça de discrétiser. Ce n'est pas pour moi le nœud du problème, même si là, on a l'impression que c'est important. C'est important, mais il y a tout de même un autre point qui est majeur.

Globalement, sans trop rentrer dans les détails techniques, ce dossier a généré beaucoup de frustration. Déjà, je remercie les rapporteurs parce que l'instruction a été longue et ça demande une certaine endurance de se remettre dedans. On a eu beaucoup d'échanges, je les remercie de nous avoir soutenus aussi pour l'instruction de ce dossier et de nous avoir tenus jusqu'au bout.

Il y a tout de même des frustrations qui ont, selon moi, été assez partagées. On a eu beaucoup de discussions autour de ça. Sur le plan clinique, quand on voit les courbes, on se dit que ça marche, sur les asymptomatiques, sur les paucisymptomatiques, mais que ça ne marche pas très bien.

Mais comme on l'a déjà dit ce matin, nous ne sommes pas tellement là pour juger sur le plan purement clinique, parce qu'il y a une commission pour ça. Mais en même temps, on garde tout de même ça en tête, on est forcément influencés par ça, puisqu'il y a tout de même tout le volet parties médicales qui rentrent en compte dans le ratio.

Il y a aussi le fait que l'on se dit qu'ils sont en analyse coût-efficacité et pas au coût-utilité, et ça leur est, à mon avis, très préjudiciable, parce que forcément, les patients, en termes de qualité de vie, y ont gagné de façon très importante, mais on ne le saura jamais, pas de façon formelle dans cette analyse. Ça génère tout de même pas mal de frustrations.

Enfin, on a des données de long terme, et ça, c'est très rare, surtout sur ce type de population, et c'est très pertinent, étant donné la pathologie. Ça fait des points où on se dit tout de même, c'est hyper intéressant, mais la question, c'est : le ratio, on y croit ou on n'y croit pas ? Un des points qui fait que l'on n'y croit pas, c'est que la comparaison est complètement naïve, c'est-à-dire qu'il n'y a aucune méthode qui permet de se dire : les groupes sont-ils comparables ?

Quand on voit le niveau d'exigence qu'on a pour un dossier cancer, où ils nous font des méta-analyses en réseau et on vient discuter si la loi était bien appliquée, et que là, il n'y a rien, on ne peut pas laisser passer ça. C'est juste impossible, en termes méthodologiques, de se dire qu'on ne sait pas qui on est en train de comparer.

On voudrait bien les croire sur parole que les patients qu'ils comparent se ressemblent, mais on n'a aucune preuve. On doit les croire sur parole et c'est tout de même un peu problématique quand on est censé évaluer le côté scientifique.

Il y a un autre point qui me dérange plus dans ce dossier que dans les autres c'est dans l'estimation des coûts. Généralement, dans beaucoup de dossiers, on se dit que le traitement est tellement cher que les autres coûts sont quasiment négligeables. Sauf que là, ces patients ont énormément de traitements, et on arrive, sur les groupes qui sont sur le *best supportive care*, à peu près 900 000 euros de coûts. Cela réduit énormément le différentiel de coûts entre le groupe *ati-cel* et le groupe *best supportive care*, sauf que ces coûts ont été estimés par avis d'experts. C'est un peu léger, d'expérience, je pense que tous les économistes ont eu affaire à ça. Ça ne veut pas dire que les experts ne sont pas de bonne foi, loin de là, mais quand on demande aux médecins de nous décrire le parcours patient et quand on a ensuite les données en vie réelle de la réelle consommation de soins, on est très surpris. C'est tout de même un point majeur qui a été estimé sur des données qui ne sont pas robustes et qui influencent largement le différentiel de coûts et l'ICER à la fin.

Sans rentrer dans des choses qui sont très techniques, la sensation que j'ai sur ce dossier, c'est que malheureusement, il y a des informations intéressantes, mais on a du mal à croire au ratio.

Peut-être qu'il est plus bas que ce qu'ils ont estimé à cause de l'utilité, mais on ne peut pas le savoir. Ce point de comparaison des groupes, pour moi, pose problème. Qu'allons-nous dire à tous les autres dossiers sur lesquels on leur parle de la loi, alors que là, on a laissé passer quelque chose où, scientifiquement, c'est compliqué à valider ?

M^{me} CHEVREUL.- Avant de laisser la parole, cela me suscite tout de même une question. Quand on est dans les dossiers de cancer on a des lois. Et si on a ces lois, c'est qu'il y en a beaucoup. Doit-on répondre exactement de la même façon quand on a des tailles d'échantillons qui sont énormes ? C'est une question à laquelle je n'ai pas de réponse, mais c'est une vraie façon de se positionner. Ici, ils sont 30 patients inclus dans l'essai et on ne peut pas faire plus parce que, de toute façon, ça reste heureusement une maladie qui n'est pas très fréquente, c'est une maladie rare. Doit-on se positionner exactement de la même façon ? C'est une question qui se pose.

Si on réagit comme ça dans le cancer, c'est parce qu'on en a la possibilité, on a de grands échantillons. Sur d'autres maladies, peut-on dire exactement la même chose et dire : non, parce que vous ne l'avez pas dit, avec 30 patients...

M^{me} MARGIER.- Ce n'est pas que ça, l'argumentaire. C'est de se dire tout de même que sur le plan scientifique, c'est comme si je prenais des dossiers aux hospices, les 30 premiers, et les 30 premiers et que je les comparais. Je me ferais dézinguer, clairement, et c'est normal.

M^{me} CHEVREUL.- Je réagissais juste par rapport au cancer.

M. SERRIER.- Le corollaire de ce que vient de dire Karine, je me permets de rebondir là-dessus, parce que là, on est dans le cadre des MTI, si je ne dis pas de bêtises, c'est pour ça qu'ils passent avec 3 patients par an. La question aussi, c'est de se dire s'il n'y a pas un seuil en dessous duquel ils ne pourront de toute façon rien faire et qui devrait être exclu de l'éligibilité. C'est une question ouverte.

M^{me} CHEVREUL.- Sauf que là, c'est le CEPS qui a demandé nommément à ce qu'on le fasse.

M. SERRIER.- C'est une vraie question, parce que si à un moment donné on se retrouve avec des cas de figure où ils ne peuvent avoir que des essais à 20 patients et qu'ils n'arriveront jamais au niveau

M^{me} MARGIER.- Ils ont tout de même fait une comparaison pour la CT. Ils n'avaient pas plus de patients.

M^{me} CHEVREUL.- Ils ont fait une comparaison sur le premier critère pour la CT.

M^{me} MARGIER.- Ça veut dire qu'il y a des choses qui sont possibles quand bien même l'effectif est faible.

M^{me} CHEVREUL.- Je vais donner la parole à Clémence, puis au chef de projet.

M^{me} THEBAUT.- C'est un dossier très compliqué à appréhender. Je ne vais pas revenir sur tout ce qui a été dit. Ce qui est complexe pour moi, et c'est une remarque plutôt interrogative, c'est attention à tout ce que l'on va dire sur la question du critère GMFC, parce qu'on est tout de même dans un modèle d'analyse coût-efficacité en années de vie gagnées. Cette question persiste pour moi : on pourra dire tout ce que l'on veut sur la faiblesse de la démonstration du résultat sur ce critère, mais dans quelle mesure ce critère impacte-t-il la mortalité ? Si nous étions en analyse coût-utilité, cela pourrait être fondamental, car ils auraient pu nous présenter des utilités peu fondées.

Je voudrais recentrer, ou plutôt reformuler ma question sur les limites de la démonstration d'effets, avec ces limites méthodologiques sur la démonstration du gain de survie. C'est plus une question qu'un rapport.

Deuxième point : j'ai passé mon temps à relire le dossier, car on a l'impression qu'il y a quelque chose qu'on ne comprend pas. Pourquoi sont-ils allés chercher ce critère secondaire, alors qu'ils avaient un critère d'efficacité significatif en principal avec une analyse de comparaison indirecte ? Il me semble qu'il y a peut-être une réponse dans le modèle du NICE. Quand on examine le modèle soumis au NICE, on constate que ce qu'ils nous ont donné est un copier-coller complet de ce qu'ils ont fourni au NICE qui, lui, était une analyse en coût-utilité. Toute la structure du modèle n'a-t-elle pas été construite pour le NICE, d'ailleurs avec l'aide du NICE, du *review group* de York ? Mais le transposer en France en analyse coût-efficacité, c'est bien compliqué pour finalement mesurer la différence de survie.

M^{me} CHEVREUL.- J'ai dû rater l'épisode, parce que c'était un peu long, je me permets de prendre la parole tout de suite. Pourquoi ont-ils été capables de faire un coût-utilité pour le NICE et pas pour nous ?

M^{me} THEBAUT. Parce que les utilités étaient inacceptables pour la France, c'est ça ?

Une cheffe de projet, pour la HAS.- Effectivement. Le laboratoire a mené une étude de vignette pour informer la qualité de vie des enfants et ils ont fait le choix au dépôt de ne pas présenter d'analyses coût-utilité.

Un chef de projet, pour la HAS.- Merci Carine, car tu as soulevé un point très intéressant. C'est un point hyper intéressant, et en 2025, pour la CEESP, il y aura plein de sujets, plein de choses intéressantes pour améliorer la situation.

Pour NICE, ils utilisent les vignettes. Dans notre guide méthodologique, lors de la mise à jour de la dernière version, la CEESP a opté à la majorité pour ne pas retenir la méthode des vignettes à

cause des biais liés à l'utilisation de cette méthode en termes de validité externe et de plusieurs validités structurelles des questionnaires. Parfois, on peut faire n'importe quoi.

Sauf que le NICE a mis quelques ajustements très subtils pour des maladies rares et ultra-rares. Ils disent que dans des cas extrêmes, on peut accepter la méthode des vignettes. Sauf que NICE, quand il accepte les méthodes des vignettes, impose dans le *process* d'évaluation un support fait par des experts NICE qui n'ont pas de conflits d'intérêts, et qui disent « faute de mieux ». Comme l'a dit Karine pour ces quatre ou ces quinze patients, le NICE dit : faute de mieux, comme on a un avis d'expert au pire des cas, et qu'on peut faire confiance à une certaine tolérance d'incertitude pour ce degré d'expert, on utilise la méthode coûts-utilité. Effectivement, Clémence, ils étaient sauvés parce qu'ils ont utilisé la méthode coûts-utilité.

Pour nous, c'est un peu le revers de la médaille. Comme on dit que la méthode des vignettes, on ne l'accepte pas, le raisonnement de l'industriel était acceptable pour dire « vous n'acceptez pas les vignettes, je ne peux pas les utiliser ».

Je reviens à la question de Karine, c'est vrai qu'on doit faire attention à ne pas dire « les médicaments rares ou ultra-rares, on doit les analyser autrement parce qu'on va créer des débouchés pour les industriels pour ne plus utiliser les méthodes scientifiques ». En revanche, probablement, si dans le futur on peut penser à comment on accepte des méthodes trop risquées, sachant que la CEESP s'implique, sachant que les groupes universitaires ou les groupes d'experts en France s'appliquent pour atténuer un peu les biais des méthodes des vignettes, on peut les accepter et éviter les problèmes d'aujourd'hui.

C'est la première précision qui est hyper importante, à mon avis, pour la CEESP et pour la méthode des vignettes.

M^{me} CHEVREUL. - J'ai une question dans le *process*. Là, on se retrouve tout de même dans une situation où on est quasiment tous d'accord, à moins que j'aie mal compris. J'ai cru comprendre que tout le monde était à peu près d'accord pour dire que le coût-efficacité, ça n'a pas beaucoup de sens.

M^{me} THEBAUT. - C'est tout de même pessimiste, et en cela, ce n'est pas si mal. Comme tu l'as dit, Jennifer, on aurait pu s'attendre à ce qu'il y ait plus de bénéfices captés. C'est en cela que c'est un choix qui n'est pas inacceptable pour nous. C'est pessimiste, mais ce n'est pas si mal. Ils font ce choix pessimiste parce que scientifiquement, c'était ce qui semblait le moins critiquable pour la CEESP, mais avec un modèle qui n'a pas été fait pour une analyse coût-efficacité et qui est inutilement compliqué pour, à la fin, montrer une différence de survie. On perd en clarté. On a l'impression de se faire balader un peu. On ne comprend pas très bien. Je parle en des termes très simples.

Mais attention, pour notre jugement : est-on en train de juger la structure du modèle complet ou l'incertitude sur le résultat coût en année de vie gagnée ? Il reste le problème des coûts, parce que c'est là que ça peut avoir un effet, le GMFC. Y a-t-il potentiellement des états qui, avec ce critère secondaire, ont pu être un peu plus ou moins courts ou plus ou moins longs ? De toute façon, on a tout de même une grosse incertitude qui surpasse ça, qui est la question de combien ça coûte de prendre en charge un enfant dans les différents stades.

L'énormité de cette incertitude est vraiment plus importante que la question de critères secondaires ou primaires. Là-dessus, malheureusement, je suis d'accord que l'avis d'experts est fragile et ils auraient pu faire un effort pour aller regarder dans les bases de données de l'assurance-maladie, par exemple, le SNDS. Mais il est possible qu'il y ait, de toute façon, une très grosse variabilité, parce que la réalité, c'est que ce sont des enfants qui sont dans un état végétatif et qu'en plus, ça doit très fortement dépendre des centres et ça dépend aussi des situations familiales. Dans ces contextes pédiatriques, il y a des moments où les enfants restent à l'hôpital parce qu'ils n'ont aucun autre endroit où aller. Dans d'autres, ils ont une famille qui peut les prendre en charge. Il y a, de toute façon, une variabilité centre, pays et familial. Mais c'est sûr que cela joue beaucoup sur le RDCR.

Le dossier, globalement, est tout de même très bien présenté sur tout le début du dossier. Ça commence à se compliquer sur la modélisation. Il y a eu un gros effort de l'industriel, que je voudrais juste souligner, après je m'arrête. Parce qu'on a suggéré en échange technique qu'on ne croyait pas à leur horizon temporel, vie entière, avec une espérance de vie. Ils ont d'eux-mêmes décidé de prendre en compte notre remarque pour changer l'analyse de référence et ils nous ont transmis toutes les analyses de sensibilité. Il y avait de la bonne volonté tout de même de l'industriel.

M^{me} CHEVREUL. - On imagine qu'ils se sont retrouvés dans la situation de la majorité d'entre eux. C'est-à-dire que le modèle est fait au global. Il est d'abord fait pour le NICE parce que c'est eux qui le prennent réellement en compte pour l'entrée sur le marché. Ensuite, on leur a dit : non pas les vignettes, donc ils se sont dit : on va faire l'efficacité. Ils n'ont pas refait tout le modèle parce que c'est beaucoup trop demandant, juste pour un passage en CEESP. L'utilisation qui en est faite est peut-être plus approximative que par le NICE sur la décision. Ils se sont retrouvés dans une situation où ils avaient un modèle fait pour l'utilité qu'ils ont casé sur l'efficacité.

Peut-on réellement leur en vouloir ? Je ne sais pas parce qu'il y a le coût de l'information. Reconstruire un modèle pour produire une information qui soit plus juste, dont ils ont une certaine incertitude sur l'utilisation et l'effet sur le prix. Ce n'est pas la question du jour, je suis

d'accord, mais cela doit se réfléchir tout de même. Peut-être que cela doit se réfléchir en se disant cela a-t-il du sens de faire de l'efficacité dans ce cas ?

Cela veut dire qu'on est aussi obligé de réfléchir nous-mêmes sur le fait que si on demande des choses spécifiques à un industriel, on sait que de toute façon, quel que soit l'échange technique, il n'est pas en mesure de le faire, par exemple, changer la structure du modèle. N'est-ce pas à nous dans ces rares cas, et ça revient à ce que disait le chef de projet, de revoir un peu la doctrine de la CEESP, et peut-être accepter les vignettes. Honnêtement, l'incertitude sur les vignettes et la possibilité de les manipuler est exactement la même que sur l'avis d'experts sur la prise en charge. C'est ce qu'on voit ici. Sur le coût de la prise en charge classique sans intervention à 900 000 euros, est-ce 1,2 million ou 600 000 ?

Après, on est tout de même obligé de se rendre compte qu'effectivement, par rapport à ce que l'on aurait fait d'habitude, ce que tu dis assez bien, c'est plutôt péjoratif dans leur sens. Considère-t-on que cela éclaire ou pas la décision et que c'est utilisable dans une négociation de prix ? C'est une autre question et c'est celle qui doit revenir à la fin.

Je vous invite à plutôt le discuter dans ce sens. Parce que quelque part, quand on se dit qu'on a une réserve majeure là-dessus ou une incertitude globale majeure, on arrivera quasiment à la même chose, à mon avis.

Je vais laisser la parole à Antoine qui essaye depuis tout à l'heure de dire quelque chose. Puis au chef de projet. Tu voulais dire quelque chose ?

M^{me} THEBAUT.- Sera-t-il possible d'avoir votre avis sur l'effet sur la mortalité, pour être sûr qu'on ne part pas dans le mur ?

Un chef de projet, pour la HAS.- J'allais répondre à ce que Clémence a dit.

M^{me} CHEVREUL.- Je voulais juste laisser la parole à Antoine, parce que depuis le temps qu'il attend, j'ai l'impression que je l'ai doublé. On lui laisse la parole, après c'est toi.

M. BENARD.- Pour ce qui est de la comparabilité entre les sujets non traités et les sujets traités, je ne sais pas si je peux m'exprimer ainsi, car je ne connais pas vraiment l'origine du groupe de comparaison, il y avait un appariement sur l'âge, c'est ça ? S'agit-il d'un appariement direct sur l'âge, un appariement individuel ?

M^{me} CHEVREUL.- Non, justement, ils ne l'ont pas fait sur l'âge.

M. BENARD.- C'est un assortiment, c'est ça ? C'est globalement les deux cohortes ont le même âge ?

Un chef de projet, pour la HAS.- Ce n'est pas via des probabilités comme le score propension.

M. BENARD.- Je sais. Par exemple, il y avait un enfant dans le groupe traité qui avait deux ans. Il a été apparié avec un enfant de deux ans dans le groupe non traité ?

Une cheffe de projet, pour la HAS.- Non, justement, ce n'est pas un appariement individuel entre les patients. Ils prennent tous les patients traités par ati-cel, ils regardent les patients non traités et ils vont sélectionner uniquement les patients qui ont un âge compris dans une fourchette qu'ils ont définie dans la cohorte traitée.

M. BENARD.- OK, c'est ce que j'appelle un assortiment de fréquence.

Une cheffe de projet, pour la HAS.- Cet appariement sur l'âge a été jugé de faible qualité côté commission de la transparence.

M^{me} CHEVREUL.- Après, on n'est pas tout à fait dans la même situation, parce que la commission de la transparence l'a fait sur le critère principal. Alors que nous sommes sur le critère secondaire. C'est assez important. Après, ont-ils assez de population, justement quand on est sur ce critère secondaire, pour appareiller exactement sur l'âge ?

Une cheffe de projet, pour la HAS.- C'est toute la question, effectivement même s'ils avaient eu la possibilité, ou démontré que cela aurait été possible ou impossible, il y a des limites liées aux données disponibles.

M. BENARD.- Juste pour compléter ma question : quand il y a un assortiment d'âge, en moyenne, les deux cohortes ont le même âge, il se peut très bien qu'il y ait des enfants qui soient plus âgés dans le groupe non traité et plus jeunes dans le groupe traité. Cela pourrait aggraver le pronostic, avec un pronostic plus grave dans le groupe non traité, parce que le pronostic est essentiellement dépendant de l'âge, j'imagine, et à quelques mois près aussi. Si on ne maîtrise pas bien ce facteur âge, cela rend les deux cohortes très peu comparables.

Un chef de projet, pour la HAS.- C'est exactement ce que la cheffe de projet vous a présenté. Il y a des différences observées sur l'âge au diagnostic, et que l'âge au diagnostic est un facteur pronostique. Il y a un ajustement indirect qui est fait par type de maladie, puisque les résultats sont présentés par sous-type. Mais c'est un aveu de l'industriel qui dit : on n'a pas fait d'appariement sur l'âge entre les deux cohortes, qui est pourtant un facteur pronostique. Mais les âges sont plus ou moins comparables, mais dans le tableau, on voit que ce n'est pas forcément le cas.

Une cheffe de projet, pour la HAS.- L'âge moyen au début des symptômes dans la cohorte historique, ils ont eu des symptômes plus tardivement.

M^{me} CHEVREUL.- C'est donc plus péjoratif pour le traitement.

Une cheffe de projet, pour la HAS.- Autrement dit, les patients traités ont eu un développement de la pathologie plus tardif que les patients non traités.

M. DOURGNON.- C'est tout à fait possible dans cette situation, ce n'est pas compliqué de faire un appariement individuel sur l'âge. Ça peut réduire un peu les cohortes. C'est la taille des cohortes qui est en jeu.

Une cheffe de projet, pour la HAS.- Ils ne nous ont pas donné de preuves que c'était impossible ou qu'ils ont testé quelque chose sur le critère secondaire, quelle que soit la méthode. On est un peu dans le flou concernant ce qui était possible ou pas.

Un chef de projet, pour la HAS.- Si je peux rebondir sur ce que Clémence avait dit sur la mortalité. Pour rappel, on a la classification du type de réponse, complet, partiel, stabilisé, non stabilisé. Ensuite, on a les multiplicateurs d'effets, c'est là qu'intervient le score GMFC. Il n'intervient qu'au moment des multiplicateurs d'effets. Les multiplicateurs d'effets s'appliquent chez qui ? Chez les patients qui sont répondeurs partiels non stabilisés.

Or, ce sur quoi on avait insisté, c'est que c'est la sous-population de patients juvéniles précoces paucisymptomatiques qui ont la proportion de patients répondeurs partiels non stabilisés la plus importante, c'est-à-dire 80 %. C'est chez ces patients qu'on va avoir les multiplicateurs d'effets qui vont avoir le plus d'impact. Donc comme le stade GMFC 6 du modèle précède l'état décès, l'effet sur la mortalité sera plus important chez les patients paucisymptomatiques à travers l'estimation des multiplicateurs d'effets qui est faite grâce au score GMFC.

Tout cela pour dire que cette incertitude autour de l'estimation des multiplicateurs d'effets grâce au score GMFC a plus d'impact et est plus d'importance chez les patients juvéniles précoces paucisymptomatiques que dans les autres populations.

M^{me} CHEVREUL.- J'ai tout de même une position, c'est de dire que, quelle que soit la réserve qu'on donne pour certains points, il existe une incertitude globale sur ce projet liée à la structure du modèle, qui n'est pas construit pour l'efficacité, liée à l'estimation des coûts qui est assez difficile par avis d'experts. Cette incertitude ne permettra pas d'éclairer le CEPS, mais quelque part, on aurait pu s'y attendre depuis le début. On est dans une situation où c'est un modèle construit pour l'utilité qui est utilisé dans l'efficacité sur une population tellement petite que c'était très compliqué.

Il y a une réserve importante ici, on peut aller sur la suite, s'il vous plaît ? On avait une autre réserve importante sur la causalité d'effets et une sur laquelle vous vous interrogez sur la réserve majeure, n'est-ce pas ? C'était sur la méthode d'estimation des probabilités de transition à travers les multiplicateurs d'effets.

Dans tous les cas, on va se retrouver à se dire si on met une réserve majeure ou une importante. On peut décider, mais pour le coup, on est tout de même en train de se dire qu'on a une incertitude globale majeure. Avant que l'on entre plus dans le détail pour faire souffrir tout le monde, je voulais savoir si c'était aussi ce que l'on pouvait en sortir. Si jamais on donne une incertitude globale majeure, on assortit l'avis d'une explication en faveur de l'industriel en disant : De toute façon, on se demande comment ils auraient pu faire autrement. C'est-à-dire que c'est une réserve majeure, donc leur résultat n'a pas de sens, ou il est certainement loin de la réalité.

Je préférerais qu'on discute sur ce point plutôt que sur le reste point à point parce que globalement, ce que l'on veut savoir, c'est ce que l'on va dire au CEPS et comment ils vont s'en servir. Initialement, ils étaient dans une négociation sans avis d'efficience. J'imagine que pour l'impact budgétaire, on doit être à peu près à la même chose, mais on va voir. On va juste leur dire : de toute façon, vous allez rester sans, parce qu'on est dans une situation où ce qui est productible, il aurait été difficile de faire mieux, mais d'être beaucoup plus près.

Cela vous convient-il ? Quelqu'un aurait-il exprimé l'avis différemment ? Voulez-vous faire des propositions ? Du côté des services, cela semble-t-il refléter ce que vous avez pressenti ?

Une cheffe de projet, pour la HAS.- À titre personnel, le fait que ce soit un problème méthodologique nous orienterait plutôt vers une réserve méthodologique sur cet aspect pour l'estimation des probabilités de transition. De plus, les efforts qu'ils ont pu faire en termes de réduction de l'impact de l'horizon temporel ont fait que tous les paramètres cliniques d'efficacité ont un impact moins élevé sur la variation du RDCR. C'est ce qui fait que le RDCR n'est pas très sensible au changement qui peut être fait en analyse de sensibilité, du moins sur les éléments qui sont testés.

M^{me} CHEVREUL.- Après, c'est tout de même ce que disait Clémence, c'est à prendre en compte. C'est-à-dire que cela va dépendre énormément du coût que l'on va associer à la prise en charge standard.

Une cheffe de projet, pour la HAS.- Effectivement, sur le coût de prise en charge de la pathologie, ils ont fait une analyse de sensibilité où ils ont fait varier de plus ou moins 20 % l'intégralité des coûts. Le RDCR varie de moins de 10 %. Cela légitime-t-il une incertitude globale majeure ?

M^{me} CHEVREUL.- C'est plusieurs facteurs.

Un chef de projet, pour la HAS.- Karine, tu as donné des idées. En fait, dans ce dossier qui est un peu spécial, on a une maladie ultra-rare, NICE a décidé du statut ultra-rare, les décisions sont différentes. Nous sommes, à titre informationnel, pour le CEPS, comme tu l'as précisé.

En évoluant, avec le recul, je dirais qu'on a plusieurs questions soulevées par Clémence, notamment ce que vous avez évoqué tout à l'heure sur la comparabilité. Il y a des doutes sur certains aspects. Nous-mêmes disons qu'ils n'ont pas fait ces études de comparabilité, mais on n'est pas sûrs qu'ils peuvent les faire avec des scores de propension, parce qu'on sait que les effectifs sont trop faibles.

J'ai tendance à dire que l'industriel ne peut pas faire l'analyse coût-utilité, n'a pas fait une conception par rapport à la structure initiale, parce que le résultat de santé principal, c'est la qualité de vie. On est dans une situation un peu difficile où on est plus dans une situation d'incertitude majeure globale qui est expliquée par le fait que nous-mêmes, même si on veut arbitrer, on ne peut pas arbitrer parce qu'il y a beaucoup d'incertitudes. Personnellement, j'avais du mal à imaginer des scores de propension très simples à une seule variable avec cinq et quinze, on ne peut pas le faire.

Si on converge vers une incertitude majeure globale liée à ces principaux paramètres de comparabilité et de difficulté de l'estimation de l'effet, cela pourrait être une option plus judicieuse que de dire réserve majeure méthodologique, alors que pour la réserve majeure méthodologique, on est dans l'incertain, on n'est pas sûr, on ne connaît pas le vrai chemin. Je suis assez ferme dans ce principe.

Si on veut dire que les analyses de sensibilité l'ont fait, déjà pour l'analyse de référence, on a des doutes sur comment on peut donner un gage de crédibilité pour des analyses de sensibilité qui ne sont pas fondées sur des choses qu'on a du mal à accepter. On baigne un peu dans l'incertitude.

La question est la suivante : va-t-on dans l'incertitude majeure globale, ou, par exemple, pour FALBHATA, certains disent on veut donner le seuil très incertain ? Le problème, c'est qu'ici, on est plus dans une incertitude de grande ampleur, elle est très profonde. Je pense que c'est peut-être même plus judicieux que de dire réserve méthodologique majeure.

M. SERRIER. - J'ai raté quelque chose ou je ne comprends pas la logique. Qu'il y ait une incertitude globale majeure, pourquoi pas ? On n'en sait rien du résultat. Pour autant, il y a un point méthodologique qui pose vraiment problème et qui justifie pour moi une réserve majeure. Je ne vois pas pourquoi on ne la met pas. Encore une fois, si on la met sur la réserve globale majeure, un dossier qui revient derrière avec une comparaison comme ça, ils vont nous dire, comme ça a été fait tout à l'heure, c'est l'argument qui nous ressort : ça a été accepté dans un autre dossier.

M^{me} CHEVREUL. - Ce n'était pas les mêmes conditions.

M. SERRIER.- Je ne comprends pas pourquoi on l'accepte. Pour le coup, il n'y a rien qui impacte plus la mortalité que l'âge, de ce que j'ai compris.

M^{me} CHEVREUL.- Pour le coup, globalement, je pense que cela ne va pas changer grand-chose au résultat. On va juste faire deux propositions, pour trancher assez rapidement. Je vois bien l'avantage et l'inconvénient des deux. Je ne sais pas, Jennifer, si jamais tu veux t'exprimer de nouveau, si tu pencherais plus pour l'un ou pour l'autre.

M^{me} MARGIER.- Je ne sais pas pour l'incertitude globale majeure, même s'il y en a une. On en a discuté, mais il faut une réserve majeure sur la comparabilité des groupes. C'est trop important pour qu'on ne le dise pas.

M^{me} CHEVREUL.- Clémence.

M^{me} THEBAUT.- Je suis un peu plus positive.

Un chef de projet, pour la HAS.- Nous voulions attirer votre attention sur une option qui existe également. Nous ne vous poussons pas dans cette direction, mais nous vous informons simplement de son existence. Il s'agit d'invalidier potentiellement les résultats uniquement sur une sous-population, celle des juvéniles précoces paucisymptomatiques. C'est sur cette population que la comparabilité indirecte a le plus de conséquences. C'est une option que nous avons envisagée et dont nous voulions vous faire part.

M^{me} CHEVREUL.- Considères-tu qu'en fait la réserve majeure, le problème méthodologique, s'adresse particulièrement à cette population ?

Un chef de projet, pour la HAS.- Oui, car c'est cette sous-population qui va se voir appliquer des probabilités de transition, 80 % de la population. Les deux autres sous-populations ne représentent que 8 à 10 % des patients qui vont se voir appliquer ces probabilités de transition. L'impact sur le résultat est mineur parce que les sous-populations qui sont stabilisées n'évoluent pas, donc ils n'auront pas les probabilités de transition. Toute la problématique qui entoure à la fois la transposabilité du score GMFM au score GMFC et la comparabilité des deux populations ne concerne que 8 à 10 % des patients qui sont non stabilisés. Alors que dans la population paucisymptomatique, c'est 80 %. Cela se traduit d'ailleurs par des analyses de sensibilité qui montrent une plus forte variabilité liée à ces paramètres.

M^{me} CHEVREUL.- Si j'ai bien compris, on dit réserve majeure pour 80 % de la population ?

Un cheffe de projet, pour la HAS.- Non, réserve majeure pour la sous-population juvénile précoce paucisymptomatique, pour laquelle 80 % des patients sont non stabilisés. Pour les deux autres sous-populations, on a entre 8 et 10 % des patients qui sont non stabilisés, l'impact est mineur.

L'impact est mineur, car ces patients n'évoluent pas, donc ne se voient pas appliquer les probabilités de transition estimées à partir du score GMFC.

M^{me} CHEVREUL.- Par contre, cela voudrait dire qu'on admet une incertitude très forte dans les autres populations et globalement majeure pour celle-là ?

Un chef de projet, pour la HAS.- Après, la classification elle-même est aussi partiellement incertaine.

M^{me} CHEVREUL.- Très honnêtement, je ne sais pas ce qu'en pense la commission, mais cela devient encore plus compliqué sur du saucissonnage de populations où il y a cinq patients, dix et cinq. Honnêtement, cela ne changera rien. Justement, cela va compliquer la lisibilité du dossier pour le CEPS et cela va compliquer notre avis.

Clémence.

M^{me} THEBAUT.- C'est compliqué. Cela pose une question sur la façon dont on conçoit ce qu'est une évaluation médico-économique. Soit on considère que l'objectif est d'estimer un vrai ratio, soit on se dit que c'est un outil d'aide à la décision publique en situation d'incertitude. Il faut le faire de façon honnête, transparente, pour pouvoir servir à donner des éléments plutôt que de décider dans le noir. Si on répond par la deuxième option, et c'est aussi en fonction de la culture scientifique de chacun, l'exercice est fait de façon honnête. On a une exploration d'incertitude, il y a eu des choix conservateurs.

M. SERRIER.- Ils n'ont pas présenté l'analyse avec une autre méthode que l'appariement naïf ?

M^{me} THEBAUT.- Il y a cette question effectivement qui reste. S'ils avaient pris le critère principal avec la comparaison indirecte, on ne se poserait pas de question. On ne comprend toujours pas très bien quelle est l'histoire de ce dossier. C'est pour cela que j'ai souligné ce point.

M^{me} CHEVREUL.- Comme tu l'as dit, ils ont fait un modèle pour les utilités, et ils n'ont pas recommencé.

M^{me} THEBAUT.- Je pense, mais même ça, on ne sait pas très bien pourquoi.

M^{me} CHEVREUL.- Pour le coup, serait-on d'accord pour se dire que, méthodologiquement, à partir du moment où ça a un impact... Pardon, Martine.

M^{me} KAROUBI.- Je voulais lancer le débat. Je suis complètement incompétente, mais pourquoi n'enlève-t-on pas la réserve majeure et on la transforme en importante ?

M^{me} CHEVREUL.- Justement, c'était ma proposition. Je remercie Martine pour cette transition. J'allais proposer, peut-être pas pour la même raison que vous, qu'effectivement, si j'ai bien compris, cette situation méthodologique et ce choix ont un impact majeur sur une partie de la

population et mineur sur le reste, donc in fine, sur l'ensemble. Cela fait donc une réserve importante sur l'ensemble du résultat. Cela ne nous empêche pas de dire qu'il y a une incertitude majeure sur l'évaluation de l'efficacité, mais ce serait une réserve importante sur le choix méthodologique qui est fait là. Importante, parce que mineure d'un côté, majeure sur une partie de la population, où c'est 80 % des cinq individus. En tout, 80 % des cinq individus majeure plus mineure sur 10 plus 19, on serait plutôt dans la réserve importante pour signifier que ce n'est pas la meilleure façon de faire.

Par contre, étant donné qu'on a mis trois réserves importantes et qu'au bout du compte, on se dit, comme dit Jennifer, qu'on n'a aucune idée de savoir si on est proche ou pas de la vraie évaluation du RDCR, on signifie une incertitude globale majeure sur le niveau du RDCR, en expliquant pourquoi et en mettant un bémol sur les choix qu'ils pouvaient faire, vu cette taille d'échantillon et cette pathologie hyper rare.

Ça correspond-il à quelque chose qui est à peu près acceptable par rapport à ce qu'on s'est dit et par rapport à la structure des données du modèle ?

Hassan.

M. SERRIER. - Mon avis personnel, non, mais chacun aura son avis là-dessus. Ce qui me perturbe aussi, c'est qu'on est sur des tout petits effectifs. Quand vous dites que ça impacte peu parce qu'ils évoluent peu, mais de combien de patients parle-t-on ? Si ça concerne deux patients sur sept, si ces deux patients sont mal appariés et ont une mauvaise évolution, ça va impacter fortement le résultat. En fait, il n'y a même pas d'analyse de sensibilité pour savoir si le choix de l'appariement impacte peu, même cette sous-population. Ça reste des suppositions parce que ça concerne peu. On aurait 300 patients et on saurait qu'il n'y en a que dix qui n'évoluent pas, OK, on peut s'attendre à ce que ça impacte peu. Là, on dit que ça impacte peu, mais ça représente combien ? Sur sept, ça représente peut-être deux, c'est énorme.

M^{me} CHEVREUL. - Tu voudrais mettre majeur aussi pour cette réserve ?

M. SERRIER. - Je pense qu'il y a un problème méthodologique majeur qui a un impact sur tout. On n'a pas d'informations, on ne sait pas ce que ça impacte. Mais après, les résultats, ça ne change rien, au final, si on met une réserve globale majeure. Mon avis, c'est qu'il y a une réserve méthodologique majeure sur ce point.

M^{me} MARGIER. - La réserve majeure porte sur la comparabilité des groupes, et non sur le score GMFC. En quoi cela ne concerne que 5 % de la population ? Les groupes de base ne sont tout simplement pas comparables.

Un chef de projet, pour la HAS.- La comparabilité des groupes est utilisée surtout pour estimer les multiplicateurs d'effets. Les deux hypothèses, que ce soit la transportabilité d'un score à l'autre ou la comparabilité des groupes, sont utilisées pour l'estimation des multiplicateurs d'effets.

M^{me} MARGIER.- Mais si ta population de base n'a pas le même âge, ça impacte tout.

Une cheffe de projet, pour la HAS.- Les probabilités de transition sont utilisées uniquement pour les groupes de patients non stabilisés. En fonction de la répartition et de la sous-population, c'est pour cela que nous avons pu dire qu'effectivement, s'il y a un problème dans l'estimation de l'efficacité, il aurait probablement plus d'impact dans une sous-population. Mais en l'occurrence, cela impacte toutes les sous-populations. C'était pour éventuellement diriger la conclusion vers un autre aspect où l'on invaliderait seulement une partie des résultats dans une sous-population, mais j'ai l'impression que ce n'est pas trop le sujet, puisque cela rajouterait de la confusion.

Dans le cadre de l'avis qui a été envoyé, nous étions partis sur une réserve majeure liée à un aspect méthodologique. Restons-nous sur cela et que l'on vote l'avis ?

M^{me} CHEVREUL.- Pour le coup, la question qui se pose, c'est met-on importante ou met-on majeure sur cette partie ? De toute façon, il faut retenir que cela ne changera pas grand-chose, puisqu'on s'est tous mis d'accord sur une incertitude globale majeure en ajoutant et en revalidant au bureau une phrase qui tient compte de la difficulté, de ce qu'ils ont fait et tout ça.

On va aller très vite, parce qu'il reste trois minutes avant que la moitié des gens partent et que l'on perde le quorum, et de toute façon, cela ne va pas changer grand-chose. Qui est pour majeur et qui est pour importante ? Je commence par réserve importante. Je ne parle pas d'incertitude globale majeure, mais de la réserve importante sur celle pour laquelle il fallait décider sur les probabilités de transition. Qui est pour majeur ? Cela me semble plié. On va voter l'avis avec réserve majeure, incertitude globale majeure et en expliquant ce qu'on s'est dit jusqu'ici sur la capacité à faire autrement, pour que ce soit regardé d'un œil non péjoratif. Par contre, on n'a pas vu l'impact budgétaire.

Un chef de projet, pour la HAS.- Karine, juste une précision, soit incertitude majeure, soit incertitude majeure globale, soit réserve méthodologique majeure.

M^{me} CHEVREUL.- Je ne vais pas préciser parce que sémantiquement j'ai un peu de mal. Cette réserve méthodologique est pour l'appariement, par contre, on s'est bien dit que ce qu'on mettait était une incertitude globale majeure. L'un empêche l'autre ?

M^{me} MINAYA FLORES, pour la HAS.- Je suis désolé d'évoquer ce point, mais je crois qu'on aurait bien aimé un GT là-dessus. Parce que là, on arrive à 17h30, on n'a pas forcément décidé. La

question se pose encore par rapport à ce que l'on vote. Il y a des gens qui sont déjà partis. Je ne sais pas ce que vous en pensez, mais ça aurait été bien de discuter avant les questions super importantes qui sont en train d'être abordées en fin de CEESP.

M^{me} CHEVREUL.- OK. J'ai un autre sujet à vous présenter pour ceux qui sont intéressés. On a eu une question du CEPS sur comment ils pouvaient réutiliser des modèles. On va refaire une réunion de la CEESP, deux heures, pour traiter ce dossier, si jamais vous considérez que vous ne voulez pas le voter maintenant. Va-t-on arriver à autre chose que cette partie ?

M^{me} SILVA.- Oui, de toute façon, tout le monde est d'accord pour majeure. Après, il y a la question de comment le r méthodologique ou pas méthodologique.

M^{me} CHEVREUL.- Pour le coup, ce que je comprends, c'est que si on met une réserve méthodologique majeure, le dossier est de toute façon invalidé. Si on met une incertitude globale majeure, on peut tout de même le transmettre avec une phrase d'explication, c'est ça ?

M. BEAUFILS, pour la HAS.- De toute façon, on transmettra le dossier dans les deux cas. Dans les deux cas, le dossier est invalidé. Classiquement, on n'a jamais mis les deux, on n'a jamais mis une réserve majeure et une incertitude globale majeure, ça fait un peu ceinture et bretelles. On peut aussi considérer qu'à partir du moment où on met une réserve méthodologique majeure, le dossier est invalidé et ne pas mettre les deux.

M^{me} CHEVREUL.- Après, dans ce cas, ça veut dire que si on met une réserve majeure sur ce point, on ne peut pas expliquer l'incertitude globale majeure par le fait qu'on ait choisi un modèle non adapté.

M. BEAUFILS, pour la HAS.- On pourra dire que c'est très incertain.

M^{me} THEBAUT.- Après, ils vont arriver en audition et ils vont nous dire : on n'est pas d'accord avec ce point. On devra dire : oui, mais il y avait tout le reste donc on est tout de même en accord global.

M^{me} CHEVREUL.- On peut faire le choix qu'on veut parce que, de toute façon, in fine, le résultat reste le même. Cela veut dire qu'on a une incertitude globale majeure : soit on le met directement sur la réserve, soit on le met en bas.

Puisque tout le monde a émis une réserve majeure plus tôt sur la partie méthodologique, on va procéder de cette façon.

M. SERRIER.- Je ne suis pas totalement d'accord. On vote pour la conclusion. La conclusion parlait du dépistage et de l'analyse de l'impact budgétaire.

M^{me} CHEVREUL.- Bien sûr, mais là, on ne l'a pas encore fait, l'analyse d'impact budgétaire. On est d'accord qu'il faut qu'on fasse cette partie.

M. SERRIER.-, Mais tu veux voter l'aspect, c'est ça ?

M^{me} CHEVREUL.- Pour le coup, on va peut-être présenter l'analyse d'impact budgétaire. On a toujours le quorum. Par contre, c'est largement insatisfaisant pour les gens qui sont partis de ne pas participer au vote. Sur l'impact budgétaire, on avait d'autres remarques, parce que de toute façon, j'imagine qu'on va dans la même direction.

Une cheffe de projet, pour la HAS.- L'impact budgétaire va être simple. On ne l'a pas invalidé, car, du fait des spécificités de l'horizon temporel réduit, on s'est dit que l'impact était a priori plus faible. On présente les résultats de l'impact budgétaire.

M^{me} CHEVREUL.- Ce qui compte, c'est la rédaction de l'avis. C'est là-dessus que Hassan aurait aussi aimé s'exprimer. On a bien compris que de toute façon, si jamais on votait l'avis, c'était sur la réserve majeure et donc sur les conclusions. Ensuite, on propose qu'il y ait une réflexion sur la réécriture de l'avis, que ce soit ensuite repassé, soit en CEEESP, soit en bureau.

Ce qu'on peut faire, c'est faire comme la dernière fois, faire recirculer la réécriture de l'avis, prendre l'avis de chacun. Ça a tout de même bien marché, la conclusion de l'avis. Prendre l'avis de chacun sur la conclusion de l'avis et voter de façon conditionnelle, voter dans un deuxième temps, comme on a fait la dernière fois. Cela vous convient-il si on procède comme ça ?

Je retiens d'aujourd'hui qu'on a discuté longtemps sur réserve importante ou pas, qu'on a choisi la réserve importante, que globalement, il n'y a quasiment pas d'effet sur l'impact budgétaire, parce que vu le prix de la chose et vu la taille de la population, il y a des réserves, mais ça ne vaut pas une réserve majeure et qu'il est donc communicable et utilisable.

Ce qu'il va falloir, c'est qu'on modifie notre avis en fonction de ce qu'on s'est dit sur justement l'incertitude qui a entraîné le pourquoi et la difficulté de ce type de dossier, et on pourra réfléchir après sur comment il faut qu'on procède.

Si vous êtes d'accord, pour l'instant, on ne vote pas, on modifie l'avis de la commission, les conclusions en ce sens, on fait un échange sur les conclusions et on revotera par correspondance suite à nos discussions. Cela vous convient ?

Je vous remercie. Bonne soirée.