

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.

Audition-Avis économique-BREYANZI dans le traitement des patients adultes atteints d'un lymphome diffus à grandes cellules B, d'un lymphome B de haut grade, d'un lymphome médiastinal primitif à grandes cellules B ou d'un lymphome folliculaire de grade 3B, en rechute dans les 12 mois suivant la fin d'une immunochimiothérapie de première ligne ou réfractaire à ce traitement de première ligne (BRISTOL-MYERS SQUIBB) – ECO-EFFI 659

M. RUSCH.- Je vous propose d'aller sur la partie audition avec, je pense, une présentation préalable. Pierre a levé la main.

M. HERTZ.- Oui, Emmanuel, je fais un déport sur la question suivante, donc je vous rejoins à 15 heures 30.

M. RUSCH.- Merci Pierre, à tout à l'heure. Les cheffes de projet vont nous rappeler les éléments du dossier BREYANZI.

Une cheffe de projet, pour la HAS.- Bonjour à tous. Avec l'autre cheffe de projet, on va reprendre la main avec l'audition de BREYANZI. Pour rappel, les rapporteurs de ce dossier, que nous remercions, sont Claire Granon et Jennifer Margier.

Pour rappeler quelques éléments de contexte. BREYANZI a eu une AMM centralisée en avril 2023 suivante dans la population revendiquée au remboursement : le traitement en deuxième ligne de quatre lymphomes non hodgkiniens qui sont soit en rechute, soit réfractaires à la première ligne. Ces quatre lymphomes sont, comme je viens de le dire, des lymphomes non hodgkiniens, qui sont agressifs. Ils sont d'évolution rapide et sont caractérisés par leur présentation symptomatique.

Jusqu'à présent, la prise en charge de ces quatre lymphomes se faisait en deuxième ligne par rapport à l'éligibilité ou non à une autogreffe de cellules souches hématopoïétiques. L'industriel revendique, avec BREYANZI, un changement de paradigme et, plutôt un traitement en fonction ou non de l'éligibilité des patients à un CAR-T pour la deuxième ligne de traitement.

Avant l'audition, on vous rappelle la conclusion de la CEESP qui a été tenue le 21 novembre sur l'analyse d'efficience, une réserve majeure, trois réserves importantes et trois réserves mineures avaient été votées. La réserve majeure a invalidé la méthode sur laquelle reposait l'analyse coût-utilité, dans la population analysée. Cette réserve majeure portait sur un ensemble de limites concernant l'estimation des scores d'utilité, remettant en cause leur robustesse en lien avec les sources d'incertitudes suivantes : le recueil des données de qualité de vie, la source de données mobilisées pour estimer le score d'utilité pour la survie post-événement, la méthode d'estimation du score d'utilité pour la survie sans événement, enfin la non-plausibilité clinique des scores d'utilité correspondant aux états de santé.

Je ne reviendrai pas sur la méthode sur laquelle reposait l'analyse coût-efficacité, puisqu'elle n'avait pas été invalidée.

Aujourd'hui, l'objet de l'audition, c'est une demande de la part de l'industriel d'une requalification de la réserve majeure en deux réserves importantes.

Une cheffe de projet, pour la HAS.- Je vais vous présenter rapidement pour rappel ce qui avait été proposé initialement par l'industriel dans le rapport technique après l'échange technique.

Plusieurs sources de données étaient proposées pour estimer les scores d'utilité. On retrouvait l'essai TRANSFORM pour documenter le score d'utilité de l'état de santé survie sans évènement ainsi que le décrétement d'utilité lié à l'administration du traitement. On retrouvait également l'essai clinique TRANSCEND NHL 001, qui est un essai clinique de phase I qui évaluait l'efficacité de liso-cel en troisième ligne de traitement. Il est ici utilisé pour documenter le score d'utilité dans l'état de santé survie post-événements. Enfin, on utilisait la littérature pour estimer les décrétements d'utilité liés aux événements indésirables.

On avait une valorisation des scores d'utilité fondés sur la matrice de pondération des préférences françaises de la version 5L de l'EQ-5D.

En termes de méthode d'estimation, on avait des scores indépendants du traitement reçu et dépendants des différents états de santé. Le choix de retenir des scores indépendants du traitement reçu était justifié par l'industriel, compte tenu du parcours de soins hospitaliers qui pouvait être similaire entre l'administration du CAR-T et la greffe, ainsi qu'une guérison possible pour ces deux traitements. Les scores d'utilité étaient également ajustés sur l'âge et le sexe des patients.

Maintenant, si on regarde un peu plus la méthode d'estimation des scores d'utilité en tant que tels, dans l'état de santé survie sans évènement, il était proposé de retenir comme score d'utilité le score d'utilité moyen au moment de la randomisation dans l'essai de TRANSFORM. Ce score aboutissait à une valeur de 0,936, ce par quoi l'industriel proposait de borner ce score d'utilité par le score de la population générale française ajustée sur l'âge et le sexe, qui était de 0,894.

Pour le score d'utilité dans l'état de santé survie post-évènement, il était proposé d'appliquer un décrétement d'utilité qui était lié à la progression de la maladie, au score d'utilité estimé dans l'état de survie sans évènement. Ce décrétement d'utilité était estimé à partir de l'essai TRANSCEND NHL 001, pour lequel on avait également des données en troisième ligne.

C'était les scores d'utilité avant les cinq premières années. Comme on avait fait une hypothèse de guérison après cinq ans, il était proposé, après cinq ans, de retenir les scores d'utilité de la population générale, puisque les patients étaient considérés comme guéris.

Enfin, nous vous proposons de revoir l'analyse critique que nous avons proposée dans le projet d'avis avant audition. Il semblait que plusieurs limites pouvaient être identifiées sur ces méthodes d'estimation des scores d'utilité, et que ces limites étaient majoritairement inhérentes à l'absence d'un recueil de qualité de vie rigoureux lors de l'essai clinique TRANSFORM. C'est un point sur lequel l'industriel reviendra probablement pendant l'audition.

Mais effectivement, cette absence d'un recueil de qualité de vie rigoureux entraîne plusieurs limites, notamment pour la méthode d'estimation du score d'utilité SSE, qui semble être une méthode peu robuste, et également non plausible d'un point de vue clinique, puisque finalement, elle ne semble pas permettre de refléter réellement la qualité de vie dans un état de santé, puisque l'on ne retient que le score estimé au moment de la randomisation des patients. De plus, cette non-plausibilité est également mise en évidence par le score d'utilité à laquelle aboutit cette méthode d'estimation, qui est finalement supérieure à celui de la population générale française, ajustée sur l'âge et le sexe.

Enfin toujours lié à l'absence d'un recueil de qualité de vie rigoureux lors de l'essai clinique, il n'était pas possible d'estimer un décrement d'utilité dans l'essai clinique TRANSFORM. C'est pourquoi il a été nécessaire pour estimer le score d'utilité dans l'état de survie post-événement, d'avoir recours à une autre source de données, ainsi qu'à une méthode d'estimation qui finalement nécessite de faire une hypothèse d'équivalence qui va considérer que la désutilité liée à la progression de la maladie en deuxième ligne est équivalente à la désutilité liée à la progression de la maladie en troisième ligne.

Enfin, d'autres limites s'ajoutaient à ces principales limites qui viennent d'être décrites, notamment le choix de considérer des scores d'utilité indépendants du traitement reçu, qui n'était justifié que d'un point de vue clinique par l'industriel. Également, le choix de considérer une valeur du score d'utilité de la population générale française au-delà de cinq ans, justifié uniquement d'un point de vue clinique. Enfin, une hétérogénéité des sources de données retenues pour estimer les scores des états de santé en SSE et SPE, qui étaient insuffisamment explorées par l'industriel.

Au regard de l'ensemble de ces éléments et de ces limites, il semblait qu'il y avait une incertitude trop forte pour refléter de manière robuste et plausible la qualité de vie des patients dans les différents états de santé du modèle. Il était donc proposé la réserve majeure suivante : La robustesse des scores d'utilité n'est pas assurée en raison des différentes sources d'incertitude quant au recueil des données de qualité de vie des patients, de la source de données mobilisée pour estimer le score d'utilité en survie post-événements, de la méthode d'estimation du score

d'utilité en survie sans évènements et de la non-plausibilité des scores d'utilité correspondant aux états de santé.

Voilà les éléments que nous souhaitions vous partager avant échange avec l'industriel. On a été très rapides.

Une cheffe de projet, pour la HAS.- Après, l'industriel ne voulait pas revenir sur plus de points pendant l'audition.

M. RUSCH.- Merci beaucoup. Les rapporteurs auraient-ils souhaité rajouter un mot ? Ce n'est pas forcément nécessaire, mais c'était juste au cas où.

M^{me} GRANON.- Non, cela a été parfaitement exposé et très clairement, donc pas de souci. C'est tout à fait ça. Merci.

M. RUSCH.- C'est super. Jennifer Margier.

M^{me} MARGIER.- Tout a été dit. Ça nous a fait tous un peu hurler. Vous vous rappelez, 50 % des questionnaires d'utilité n'étaient pas remplis à l'inclusion. Autour de nous, je pense qu'on fait tous, ou on connaît tous des gens qui font des études. S'il y a bien un questionnaire qui est « facile » d'avoir, c'est tout de même celui de l'inclusion. Cela en dit tout de même long sur le peu de sérieux qu'ils ont eu à récolter cette utilité. C'est tout de même bien dommage. Je ne sais pas ce qu'ils vont nous dire comme arguments. Ils avaient avancé que le Covid avait rendu les choses difficiles. Mais ils avaient d'autres questions à poser à l'inclusion, et ils les ont posées tout de même. Je ne sais pas ce qu'ils vont nous dire, mais je suis curieuse de savoir ce qu'ils vont avancer comme argument.

M. RUSCH.- Très bien. Re-bonjour Hassan. Savez-vous si les représentants de Bristol-Myers Squibb sont présents dans l'autre salle d'attente ?

M. BEAUFILS, pour la HAS.- Ce n'est pas sûr, mais on peut changer de réunion.

M. RUSCH.- Ils étaient programmés pour quelle heure exactement, Alexandre ?

M. BEAUFILS, pour la HAS.- 14 heures 45, dans dix minutes. Si vous avez d'autres choses à dire.

M. SERRIER.- Si je peux me permettre de prendre la parole parce que je ne vois pas de main levée. Il me semble, peut-être que je confonds, était-ce bien sur ce dossier où il y avait un argument qui était que ça ne changeait pas grand-chose, qu'il n'y avait pas d'impact fort sur l'utilité ? Est-ce ça ou pas ?

Une cheffe de projet, pour la HAS.- C'est un élément sur lequel ils vont revenir effectivement pendant l'audition, puisqu'ils l'ont proposé dans leurs diapositives et dans leurs observations écrites.

Après, on a des analyses de sensibilités. C'est un point dont on souhaitait vous parler après l'audition, mais on peut en parler maintenant. On a trois analyses de sensibilité en scénario, mais qui ne permettent pas complètement d'explorer l'incertitude autour des points sur lesquels nous avons identifié qu'il y avait de l'incertitude. Effectivement, il y a des analyses de sensibilité déterministes univariées qui ont été conduites et qui mettent en évidence un faible impact des paramètres qui sont retenus sur le RDCR. Maintenant, c'est un point qui n'enlève rien à l'incertitude autour de la méthodologie adoptée pour l'estimation des scores d'utilité. Effectivement, c'est un point sur lequel l'industriel reviendra.

M. SERRIER.- OK Merci.

M. RUSCH.- D'autres remarques ? D'autres questionnements ?

Un chef de projet, pour la HAS.- Ils ne reviennent pas sur l'incertitude (*inaudible 1.10.20*). Ils ne donnent pas une réponse par rapport à ça.

M. RUSCH.- On reverra ça peut-être après l'audition. Je vous propose de migrer de ce fait, de quitter cette réunion pour aller sur l'autre lien. A tout de suite.

(Sébastien Branchoux, Florian Colrat, François-Emery Cotté et Gérard de Pouvourville rejoignent la séance.)

M. RUSCH.- Bonjour, du côté de Bristol Meyer, êtes-vous tous présents ?

M. COTTE.- Oui, on est tous là.

M. RUSCH.- Très bien. Je vais vous proposer de commencer cette présentation dans le cadre de cette audition auprès de la Commission d'évaluation économique et de santé publique. Je suis moi-même, Emmanuel Rusch, vice-président de la commission. Hassan Serrier est également là, vice-président de la commission et tous les autres membres de la commission sont présents pour vous écouter. Nous vous laissons en même temps vous présenter.

M. COTTE.- Bonjour à tous et merci de nous recevoir aujourd'hui pour échanger sur notre médicament BREYANZI. Je suis François-Emery Cotté, directeur HEOR pour Bristol-Myers Squibb. Oui, je suis accompagné de Sébastien Branchoux qui est biostatisticien, qui fera l'essentiel de notre présentation aujourd'hui, et de Florian Colrat, qui est en ligne, également pour BMS. Le professeur Gérard de Pouvourville est également présent et pourra nous apporter son éclairage au sujet du principal point de cette audition, que sont les données d'utilité. En effet, Gérard de Pouvourville nous a accompagnés sur ce travail de modélisation, ainsi qu'un clinicien, le Professeur Catherine Thieblemont, chef de service d'oncohématologie à Saint Louis.

Quelques éléments de contexte rapidement. Les lymphomes non hodgkiniens sont un groupe hétérogène de maladies diffuses de prolifération anormale des cellules lymphoïdes, le plus souvent issues de la lignée B. Les lymphomes diffus à grande cellule B sont le type le plus important le plus courant des lymphomes non hodgkiniens et représentent 30 à 40 % de ces lymphomes. A noter que les cliniciens considèrent qu'il existe une possibilité de guérison à chaque stade de la maladie.

Pour ce qui est de BREYANZI ou liso-cel, c'est un CAR-T qui dispose d'une AMM dans les lymphomes en troisième ligne et plus, qui était porté par l'étude TRANSCEND NHL 001, mais pour lequel nous n'avons pas demandé de remboursement en France. Il dispose également d'une AMM en deuxième ligne, portée par l'étude TRANSFORM, qui a permis d'obtenir un AP et une ASMR III en septembre dernier. Il s'agit donc d'une primo-inscription. Le dossier médico-économique a été déposé en mai et l'échange technique a eu lieu en septembre dernier.

Pour ce qui est de notre sujet aujourd'hui, nous avons reçu le projet d'avis pour BREYANZI, qui recense au total trois réserves mineures, trois réserves importantes et une majeure. Ce projet d'avis indique tout d'abord que l'analyse coût-efficacité en années de vie gagnées est valide, c'est-à-dire qu'aucune limite majeure n'est susceptible d'invalidiser le RDCR estimé. En revanche, il considère une réserve majeure sur la mesure et la valorisation des états de santé en termes d'utilité, compte tenu du recueil et des sources de données, de la méthode d'estimation proposée et de la non-plausibilité des scores d'utilité. C'est donc sur ce point relatif au score d'utilité que nous vous proposons de reconsidérer votre avis et de requalifier cette réserve majeure en deux réserves importantes. J'y reviendrai en conclusion.

Je passe maintenant la main à Sébastien, qui vous apportera des éléments de clarification sur chacun des points de la réserve majeure.

M. BRANCHOUX. - Merci François-Emery. Nous souhaitons tout d'abord apporter une clarification quant au recueil des données de qualité de vie dans l'essai TRANSFORM. En effet, dans le projet d'avis économique, il est noté qu'un recueil rigoureux de la qualité de vie n'a pas été réalisé lors de la phase de randomisation. La méthodologie de recueil des données qualité de vie prévue au protocole de l'étude était appropriée, notamment le rythme de recueil. Malheureusement, la crise sanitaire liée à la Covid-19 a impacté le recueil des données de qualité de vie au cours de l'essai, mais aussi lors de la phase de randomisation.

Vous voyez ici, sur le haut de la diapositive, un graphique qui représente la période de randomisation de l'étude internationale TRANSFORM en marron clair, et également la phase de pandémie mondiale en gris clair. L'histogramme représente la proportion cumulée de patients randomisés au cours du temps. Comme vous pouvez le voir sur le graphique, on note une

concomitance des deux périodes avec plus de 60 % des patients qui ont été randomisés durant la phase de Covid. Compte tenu du chevauchement des périodes, les investigateurs ont rencontré des difficultés particulières. Le recueil des données sur tablette électronique a été impacté par les mesures sanitaires mises en place.

En effet, avec l'instauration des gestes barrières, des difficultés importantes ont été rencontrées pour remplir les auto-questionnaires sur tablette lors des visites en présentiel. Il y a également des visites en distanciel qui ont été mises en place et durant ces visites, les patients ne pouvaient pas remplir le questionnaire.

La crise sanitaire a donc impacté le recueil des données de qualité de vie, et ce, dès l'inclusion des patients. Cette situation était inédite. Nous avons d'ailleurs sensibilisé le SEM Éco à cette problématique lors de la rencontre pré-dépôt en décembre 2022.

La qualité de vie est un élément important dans la prise en charge des patients atteints de lymphome. Nous avons souhaité réaliser l'exercice complet d'évaluation médico-économique en réalisant une analyse coût-utilité et également une analyse coût-efficacité, malgré les taux de complétion relativement faibles des questionnaires à la randomisation.

Je passe maintenant aux valeurs d'utilité dans l'état survie sans évènement. Nous souhaitons désormais apporter une précision sur l'approche retenue pour estimer le score d'utilité dans l'état sans évènement, cette approche ayant motivé en partie les réserves méthodologiques majeures. Je vous ai décrit dans la diapositive précédente le recueil des données qualité de vie au cours de l'essai. Les taux de complétion étaient à peu près de 50 % à la randomisation. Il était inférieur à 50 % dès la première visite de suivi, puis diminuait au cours du temps.

Le projet d'avis critique le fait d'avoir considéré des scores d'utilité indépendants des traitements reçus dans l'état sans évènement, en apportant uniquement une justification clinique au regard du parcours de soins similaire entre les patients traités par le bras standard, à savoir la greffe, ou le CAR-T, liso-cel.

Compte tenu des données manquantes, il ne nous a pas semblé pertinent, d'un point de vue méthodologique statistique, de réaliser un test pour mettre en évidence une différence à la randomisation entre les deux bras, puisque l'information était disponible uniquement pour 50 % des patients. Dans ce cas, un biais de non-réponse n'est pas à exclure.

Par ailleurs, avec le nombre de patients inclus dans TRANSFORM, il n'a pas été possible de réaliser les imputations des données manquantes. C'est pour cela que l'on a uniquement apporté une justification clinique, mais bien détaillée.

Si je passe aux valeurs d'utilité qui étaient élevées, c'est vrai elles l'étaient, mais est-ce l'espoir de guérison qui a causé ces valeurs ? Cela reste encore à investiguer. Néanmoins, il semble que ce soit fréquemment le cas dans les essais cliniques menés auprès de patients traités par CAR-T, par exemple dans l'essai TRANSCEND NHL 001, un essai BMS chez les patients traités par CAR-T en troisième ligne et plus.

Par ailleurs, dans l'étude TRANSFORM, les patients du bras standard pouvaient bénéficier de liso-cel dans le cadre du *cross-over*. Cet espoir de guérison pouvait être présent dans les deux bras et peut expliquer ces valeurs d'utilité similaires et élevées par rapport à la population générale.

Enfin, nous souhaitons rappeler l'approche conservatrice que nous avons retenue pour estimer la valeur d'utilité dans l'état sans événement, en accord avec une des questions de l'échange technique suggérant d'ajuster les valeurs d'utilité lorsque ces dernières étaient supérieures à celles de la population générale. C'est ce que l'on a cherché à illustrer en bas de la diapositive.

Pour les deux bras, la valeur d'utilité moyenne à la randomisation lorsqu'elle était supérieure à la population générale, a été bornée par la valeur d'utilité en population générale. A cette valeur, on y a ajouté des décrets reflétant le profil de tolérance de chaque traitement, les événements indésirables, également un choix conservateur, comme décrit dans l'avis économique, puisque l'on a considéré un décrement d'utilité lié aux chimiothérapies avant l'administration de liso-cel.

Il en résulte que les valeurs d'utilité dans chaque bras sont différentes, en accord avec le profil de tolérance de chaque stratégie thérapeutique. Après cinq ans, en accord avec la possibilité de guérison des patients, les patients retrouvaient la qualité de vie de la population générale dans chaque bras.

En conclusion, ce qu'il faut retenir de cette diapositive, c'est que nous n'avons pas pu tester la différence entre les deux traitements dans l'essai TRANSFORM. En revanche, les valeurs d'utilité pour chaque bras, implémentées d'un modèle, étaient bien différentes.

Si je passe maintenant aux valeurs d'utilité dans l'état post-événement, dans le libellé de la réserve majeure, la source de données mobilisées pour estimer ce score d'utilité en survie post-événement a été critiquée. Je vous ai présenté dans les deux diapositives précédentes les taux de complétion et le problème de recueil de données dû à la pandémie dans l'essai TRANSFORM. A la date de l'échange technique, nous n'avons pas identifié de valeur d'utilité française dans la même indication.

Nous avons donc privilégié l'essai TRANSCEND BMS NHL 001, dans lequel, je le rappelle, des patients ont été traités par notre traitement liso-cel, en troisième ligne et plus. Puisqu'on avait

accès aux données individuelles patients, on a retenu uniquement les patients en troisième ligne pure et non en troisième ligne et plus, afin de se rapprocher au maximum des caractéristiques des patients inclus dans TRANSFORM. Dans cette étude, le décrétement d'utilité a été estimé à -0,079. Cette valeur de décrétement semble cohérente au regard des données disponibles dans la littérature. On a regardé dans l'avis d'efficience de KYMRIA en 2019, dans la même indication en troisième ligne et plus, et ce décrétement a été estimé à -0,12 chez les patients dans des lignes un peu plus avancées parce qu'ils étaient en troisième ligne et plus.

Nous avons d'ailleurs réalisé et soumis une analyse en considérant ce décrétement d'utilité. L'impact sur le RDCR était limité, puisque le RDCR variait de -3 %. Cette diminution du RDCR illustre également le côté conservateur de l'analyse de référence.

Enfin, en lien avec la guérison des patients à cinq ans post-événement, nous rappelons que l'impact de ce décrétement est limité dans le temps, parce que l'horizon du modèle était de 20 ans et après 8 ans, les patients pouvaient guérir.

Enfin, nous souhaitons rappeler aujourd'hui toutes les analyses de sensibilité qu'on a pu réaliser pour apprécier l'incertitude autour du RDCR exprimé en euros par QALY, puisqu'il est noté dans l'avis d'efficience que l'industriel n'a pas suffisamment exploré cette incertitude. Vous pouvez voir ici sur la droite de la diapositive, le tableau complet d'analyse de sensibilité déterministe que nous avons réalisé. Il a été versé au dossier en annexe 6. Par formalisme, nous avons discuté l'impact uniquement des dix premières variables que vous retrouvez en haut du tableau. Ces variables ont été discutées dans le rapport technique et sous la forme d'un diagramme en tornade.

Dans ce tableau, on retrouve les variables ordonnées par ordre décroissant d'impact sur le RDCR. On y voit que les variables d'utilité ici illustrées en jaune étaient en position numéro 20, 24 et 26 et que leur impact était au maximum de 2 % sur le RDCR.

On a également réalisé des analyses en scénario, au nombre de trois, pour apprécier l'incertitude autour des valeurs d'utilités. La première, celle que je vous ai décrite préalablement, en prenant le décrétement de Juliette. On a pris -0,12 à la place de -0,079 et l'impact était de -3 %. Nous avons également réalisé une analyse en scénarios, en considérant l'EQ-5D-3L, l'impact était de 7 %. Enfin, on a réalisé une analyse sans prendre en compte l'ajustement des valeurs d'utilité au regard du vieillissement de la cohorte dans le modèle. L'impact était ici marginal.

Nous considérons donc que l'incertitude entourant le RDCR exprimé en euros par QALY a été explorée dans le dossier et qu'elle est limitée. Nous comprenons qu'il aurait fallu intégrer davantage de variables dans le diagramme en tornade, et pas seulement les dix premières. Mais

au regard de l'impact limité sur les valeurs d'utilité, nous avons présenté dans l'analyse de sensibilité complète toutes ces variables.

François-Emery, je te laisse conclure.

M. COTTE.- Merci Sébastien. Comme vous l'avez compris, ce travail de modélisation a principalement souffert d'un problème de recueil des données de qualité de vie pendant l'essai TRANSFORM. Il s'agissait d'une situation exceptionnelle. Nous avons identifié cette difficulté très en amont et avons cherché avec notre comité scientifique la méthode la plus adaptée et la plus acceptable pour surmonter cette limite, comme vous l'a présenté Sébastien, notamment en utilisant les données disponibles, par exemple celles de l'autre essai de BREYANZI dans la même maladie, l'essai TRANSFORM.

Il nous semble que la méthodologie proposée concernant les scores d'utilité était la plus correcte possible et que l'exploration de l'incertitude soumise au dossier a montré un impact limité de ces variables, répondant ainsi à la définition d'une réserve importante.

Par ailleurs, les valeurs d'utilité que nous avons estimées nous paraissent plausibles, ou tout au moins raisonnables au vu de leur cohérence avec celle du récent avis de YESCARTA dans la même indication. En effet, comme vous le voyez, sans retenir l'impact des décrets, la valeur d'utilité utilisée dans notre modèle pour l'état sans événement présente un intervalle de confiance similaire à celui d'évaluation de YESCARTA.

C'est pour toutes ces raisons que nous souhaitons que cet avis soit compris et utilisable et que nous vous proposons de modifier la réserve majeure en deux réserves importantes qui permettront de souligner à la fois les limites de notre analyse coût-utilité, que nous ne contestons pas, sans pour autant l'invalidier totalement. Une première réserve importante pourrait concerner les conditions de recueil des données, et la seconde le manque de validité externe de la valeur estimée pour l'état post-événement.

Je vous remercie de nous avoir écoutés et nous sommes prêts à répondre à vos questions si vous en avez.

M. RUSCH.- Merci pour cette présentation. Effectivement, on avait été attentifs à la question de la qualité du recueil des données et on avait bien pris conscience et intégré la question de la temporalité. Il n'empêche qu'effectivement, ça induit un recueil parcellaire sur ces indicateurs. Je regarde s'il y a de la part des membres de la CEESP, des demandes de prise de parole ou des questionnements. Je vois que Marc Bardou a levé la main. Je lui cède la parole.

M. BARDOU.- Bonjour, c'est peut-être une question qui est un détail. Vous nous avez expliqué le fait qu'il y a eu 50 % des questionnaires de qualité de vie qui étaient remplis à l'inclusion était lié

à l'impact de la pandémie, au fait que les essais de recrutement étaient pendant la pandémie. Savez-vous, chez les patients recrutés avant la pandémie, quel était le taux de remplissage des questionnaires de qualité de vie ?

M. COTTE.- Je ne l'ai pas.

M. BRANCHOUX.- C'est une bonne question.

M. COTTE.- Le taux de remplissage de ceux qui étaient avant la pandémie.

M. de POUVOURVILLE.- C'est dans le rapport et c'est de mémoire moins de 50 %. C'est dans le rapport.

M. BRANCHOUX.- Ça, c'est au total. Mais avant le début de la pandémie ?

M. COTTE.- C'est une bonne question. Je ne suis pas sûr qu'on ait la réponse là tout de suite. Mais ce qu'on sait, c'est que dans nos autres essais qui étaient en dehors de la pandémie, notamment l'essai TRANSCEND, on avait tout de même 100 % de remplissage à la randomisation pour l'essai TRANSCEND qui s'est passé en 2007-2017. A priori, on était dans ces eaux-là.

M. de POUVOURVILLE.- Comme j'ai lu le rapport récemment, à ma connaissance, cette information est dans le rapport.

M. COTTE.- Gérard nous dit que l'information dans le rapport et on va essayer de la trouver. Florian, si tu peux regarder.

M. BARDOU.- Merci beaucoup.

M. RUSCH.- On regardera également si nécessaire. Y a-t-il d'autres prises de parole, questionnements ? Hassan.

M. SERRIER.- Je vais peut-être dire un mot, revenir sur ce que vous avez présenté. Merci pour la présentation. Nous avons bien compris que le Covid avait posé des problèmes de recueil. Les arguments que vous nous avancez, on les avait bien en tête. La difficulté, c'est que premièrement, sur les difficultés Covid, on connaît un peu le sujet. Personnellement, je connais un peu le sujet. Dans le cas des essais, j'ai compris un peu mieux avec vos histoires de tablettes la difficulté, mais je suis malgré tout surpris du faible taux de remplissage. Parce que dans des études que nous avons menées sur la période Covid, pour les données comme ça, il y avait plutôt des bons taux de remplissage, surtout qu'il y avait des possibilités alternatives, j'imagine, à mettre en place pour le faire.

Mais si on admet que le Covid a impacté votre étude, malheureusement, ça ne change rien au fond du problème. C'est-à-dire que ce soit de la faute du Covid, d'un élément extérieur ou autre, on se retrouve avec des données qui sont de mauvaise qualité et des choix derrière. Là, c'est un

peu aussi, sans rentrer dans le détail de ce que vous proposez, même si ce que vous proposez, c'est le meilleur qu'on puisse faire dans le détail, ça reste insuffisant dans le sens où c'est entaché de biais importants.

Je pense que c'était important de le remettre dans le contexte. On pointe du doigt un recueil de données qui a été biaisé par votre faute ou non, là en l'occurrence, a priori, c'est le Covid, mais le résultat est le même. Si l'étude avait été complètement arrêtée dans le cadre du Covid, vous n'auriez pas de résultat, et pour autant, vous n'auriez pas pu sortir quelque chose du tout. Là, vous avez essayé, c'est très bien. Malheureusement, c'est entaché de biais importants. C'est un peu ça l'idée que je voulais remettre dans le contexte.

M. BRANCHOUX. - Je ne sais pas si c'est vraiment une notion de biais, parce que là, la réserve était liée à l'incertitude.

M. SERRIER. - Excusez-moi, si le terme n'était pas bon, ce que je voulais dire, c'est que les résultats étaient entachés d'une forte incertitude, oui, si vous voulez qu'on le prenne comme ça.

M. BRANCHOUX. - C'est vrai qu'on n'avait pas pu avoir de questionnaire papier dans le cadre de l'étude, puisque le protocole a été mis en place avant la pandémie. Peut-être que s'il y avait eu un auto-questionnaire papier qui avait pu être mis à disposition pour recueillir d'autres valeurs d'utilité pour les patients, cela aurait été le bon augure. Cependant, ce n'était pas prévu au protocole.

Après, je pense que l'incertitude, on a démontré qu'elle était tout de même assez faible, notamment au regard des cinq ans post-événement où les patients ont retrouvé la qualité de vie des patients en population générale. Je pense que l'incertitude est plutôt limitée et certes, nos valeurs n'étaient pas forcément celles issues de l'essai clinique. Néanmoins, elles reflétaient celles qu'on a pu observer récemment dans l'avis d'axi-cel. Elles sont très proches au centième, voire au millième. Les intervalles de confiance étaient sensiblement les mêmes. On aurait attendu, on aurait bien aimé faire un scénario avec ces valeurs, mais je pense que l'incertitude aurait été similaire dans notre modèle médico-économique.

M. SERRIER. - Merci.

M. de POUVOURVILLE. - Bonjour Hassan. Je comprends bien ce que vous venez de développer. A l'évidence, quand on a un essai clinique sur lequel on a des valeurs d'utilité, on est bien contents de pouvoir les utiliser. L'appréciation qu'on peut avoir de ce problème, je la mettrai peut-être sur un plan un peu différent, qui est : y a-t-il assez d'éléments dans le rapport pour estimer que les valeurs d'utilité qui ont été utilisées sont plausibles ? François-Emery a donné quelques éléments pour justifier de cette possibilité, c'est le premier point.

Le deuxième point, le biais est le même dans les deux bras.

Le troisième point, les valeurs choisies sont finalement probablement conservatrices par rapport à ce que l'on a pu observer dans d'autres avis d'efficience, dans des modèles économiques des CAR-T cells.

Le quatrième point, si on s'intéresse finalement à une perspective au fonds de pertinence de l'importance de cette incertitude par rapport à la décision, on s'aperçoit que l'incertitude majeure sur le RDCR n'est pas dans les données d'utilité. Elle est dans les données d'efficacité. Ça, Tornado l'a bien montré.

Je me pose donc la question dans une perspective de : 1 : peut-on considérer tout de même qu'il y a des éléments qui confortent la plausibilité ? 2 : le biais est le même dans les deux bras. 3 : ce sont des valeurs plutôt conservatrices et 4 : finalement, si l'on réfléchit en termes d'impact au fond de cette incertitude comme information à donner au payeur, le facteur incertitude principal, ce ne sont pas les utilités, c'est ce que démontrent les analyses de sensibilité, que les valeurs choisies soient possibles ou pas, d'une certaine façon. Parce que là, c'est une analyse de sensibilité déterministe avec des bornes qui sont plausibles puisqu'elles ont été constatées par notre rapport.

Je crois que le fond du problème, c'est un peu ça. C'est de se dire n'est-ce vraiment pas plausible ? Sachant qu'on peut évidemment regretter qu'il n'y ait pas et se dire c'est une réserve majeure parce qu'il n'y a pas les données qu'on veut. Mais d'une certaine façon, les choix qui ont été faits sont-ils si critiquables que ça ? C'est un peu ce que je voulais amener comme élément de débat dans cette discussion. Merci beaucoup.

M. RUSCH.- Merci pour cette approche des arguments. Un chef de projet veut prendre la parole.

Un chef de projet, pour la HAS.- Bonjour. Je remercie l'industriel et son consultant pour ces éléments. Juste, ce n'est pas pour demander des réponses ou pour ouvrir un débat, juste des éléments par rapport à ce qui a été dit par les membres de la CEESP.

Là, on a des éléments où il y a des données qui manquent, comme vous l'avez expliqué c'est le Covid. Mais bon, c'est un fait. On a donc des données manquantes, un pourcentage de données manquantes très important. Vraisemblablement, il y a un biais, où il n'y avait pas d'analyse pour dire que cela touche les deux en même temps. Et même si cela touche les deux au même temps, on n'en sait rien au niveau de l'effet observé. Il y a une incertitude qui est liée vraisemblablement un potentiel biais. On ne peut pas dire biais, parce que vous n'avez pas fait cette analyse et vous ne pouvez pas la faire puisqu'il y avait une attrition très importante de manière longitudinale. Après, les questions qui ont été pointées dans cette réserve, c'est qu'il y a un ensemble

d'éléments, un faisceau d'éléments séquentiels qui commence par le problème du recueil de données et qui sont très liés.

On aurait pu avoir des réserves importantes qui sont en quelque sorte indépendantes. On aurait pu, et on a opté pour son choix. La preuve, c'est qu'historiquement, la HAS a opté pour ce choix, pour les produits de Covid eux-mêmes. Si on vous invite à voir les produits Covid, vous allez voir qu'on n'a pas donné des réserves majeures sur le Covid du fait du contexte. Et il y avait d'autres éléments qui pouvaient pallier en quelque sorte. Je pense qu'il y a tout un faisceau qui est vraiment interdépendant. Vous avez pris des morceaux de certaines parties et qui revient toujours au premier élément.

Après, la non-plausibilité est claire, si on ne peut pas spéculer sur... On peut bien sûr spéculer, c'est tout à fait légitime, on est dans un travail scientifique exploratoire, ou une publication. On peut dire s'il y a de l'espoir. On parle aujourd'hui de l'espoir, de l'innovation, tout, mais encore faut-il que cet espoir soit démontré sur une bonne base.

La chose qui nous handicape et qui est très claire, aussi bien pour le payeur que pour l'évaluateur que pour le professionnel de santé lambda, c'est qu'après, quand on lit des choses comme ça, il y a quelque chose du puzzle qui manque.

Après, si l'on cherche des analyses de sensibilité pour dire : on a la chose qui manque par rapport à l'essentiel, par rapport à l'essai pivotal. Du coup, la non-plausibilité s'additionne, émerge avec le problème de l'incertitude et crée toute cette incertitude sur cet ensemble de faisceaux.

Je reviens à l'efficacité. Il y a des éléments sur l'efficacité qui ont été mentionnés, mais je ne suis pas complètement d'accord. Parce que notre analyse sur l'efficacité est en ligne et cohérente avec celle de l'avis de la CT. L'enjeu ici se pose au niveau de la collecte des données de vie avec les problèmes de méthodes d'estimation. On ne vous demande pas de faire les méthodes d'imputation multiples, *Carried forward*, de la *best carried all of worst*, on le sait, mais c'est une conséquence malheureusement qui s'impose. C'est pour cette raison qu'on ne l'a pas mis qu'au début.

Pour nous, je rejoins à peu près, dans la discussion, le point de vue de Hassan. Le vrai problème, c'est que cela n'ajoute rien le Covid ou les autres approches. Vous allez faire certaines analyses de sensibilité, mais ces analyses de sensibilité manquent d'ancrage par rapport à leur robustesse initiale. On n'a pas de repères pour dire qu'une analyse de sensibilité ait un sens. Cela n'enlève rien de la quantité de travail que vous avez fourni et la reconnaissance du résultat de santé en termes d'années de vie gagnées. Je vous remercie.

M. de POUVOURVILLE.- Merci et bonjour. Je comprends très bien ce que tu dis. Mais en même temps, au fond, le vrai cœur de ton argument, c'est que les éléments apportés pour justifier la plausibilité des estimations qui ont été intégrés dans le modèle ne te satisfont pas, en l'absence d'un point d'ancrage. C'est un peu ça, ce que tu dis. Tu dis au fond, ça ne sert à rien d'aller chercher les données ailleurs parce qu'il n'y a pas le point d'ancrage. Finalement, on ne peut pas, on a une incertitude majeure liée à l'ignorance. Je ne porte pas jugement de valeur, je te dis simplement ça.

Mais cette question se pose très souvent finalement dans les exercices qu'on fait. Par exemple, quand on fait une modélisation Markov et qu'on a des états de santé pour lesquels on n'a pas de valeur d'utilité recueillie, par exemple en France, et qu'on va chercher des données de la littérature, c'est un peu ce qu'on a fait. Dans ces cas, ne faudrait-il pas d'une certaine façon disqualifier toutes les études médico-économiques qui sont faites, dans la mesure où les données de qualité de vie qui sont utilisées en France dans les modèles de Markov sont généralement issues de la littérature, pas vérifiables, pas contrôlables, donc avec une plausibilité très négligeable.

Je me demande s'il n'y a pas un peu deux poids, deux mesures. Je vais loin là, mais tu connais mon franc-parler. Je vais un peu loin. L'idée c'est : y a-t-il véritablement une absence totale de plausibilité des hypothèses qui ont été faites par BMS dans ce modèle ?

Après, j'ai dit mon dernier mot, c'est évidemment à la CEESP de délibérer et de juger de la pertinence de ces remarques.

En tout cas, merci de cette clarification.

Un chef de projet, pour la HAS.- Je voudrais ajouter quelque chose aussi par rapport aux données qui ont été fournies. Il y avait des données qui n'ont pas fait l'objet d'un dépôt préalable, celles portant sur axi-cel. Ce sont de nouvelles données qui ont été additionnées, et par rapport à la procédure, qui ne faisaient pas l'objet, comme vous le savez, d'une manière équitable entre tous les industriels, de comment traiter les sujets.

En abrégé, parce qu'après, c'est la CEESP qui est souveraine pour discuter par la suite, je permets de répondre à Gérard. L'accès à la revue de littérature, comme vous le savez Gérard, c'est un dernier recours, quand il est justifié et quand il y a un ensemble d'éléments le justifiant. Ici, ce n'est pas le cas. Ici, il y avait des éléments qui manquaient, il y avait des éléments de méthode liés à la méthode d'estimation, il faut regarder le rapport. Puis il y avait des éléments de plausibilité qui sont identifiés au niveau de l'estimation dont nous spéculons. Est-ce bon ? N'est-ce pas bon ? Il y avait les éléments de la littérature qui peuvent conforter parfois, mais ce ne sont pas des éléments dont la plausibilité en tant qu'analyse de sensibilité qui permettent de pallier le

handicap, les limites, qui est lié aux données, aux problèmes de la méthode et de la plausibilité. Merci.

M. RUSCH.- Merci à vous. Nous sommes arrivés au bout du temps de notre audition. Merci pour la qualité de l'échange. C'est toujours important, y compris pour la CEESP d'avoir ces temps d'échange. Je vous remercie et vous souhaite une bonne journée. Merci. J'incite mes collègues à nous rejoindre sur l'autre lien.

(Sébastien Branchoux, Florian Colrat, François-Emerly Cotté et Gérard de Pourvoirville quittent la séance.)

M. RUSCH.- Tout le monde est revenu. Merci pour ces interventions au cours de cette audition, pour ceux qui sont intervenus. Cela suscite-t-il des commentaires ? J'avais juste un commentaire. J'avais été un peu surpris de la mise en avant de l'espoir de guérison, mais je n'ai pas reposé la question. C'était en lien avec la plausibilité clinique, mais bon je ferme la parenthèse. Comme nous avons été au bout de notre temps, je n'ai pas reposé cette question. Je trouvais ça un peu curieux, mais pourquoi pas, c'est une hypothèse comme une autre. De la part des membres de la CEESP, cela a-t-il changé votre perception ? Claire.

M^{me} GRANON.- Je rebondis sur ce que tu viens de dire. Si c'était si important pour les patients, cet espoir de guérison, c'était une raison supplémentaire pour recueillir cet espoir, et avoir des questionnaires de qualité de vie. Je n'ai pas du tout été convaincue par les difficultés techniques, surtout quand on nous explique qu'on n'a pas pu faire des questionnaires papier. Ça me semble un peu limite comme explication.

Je n'ai pas du tout été convaincue non plus sur le fond de la stabilité des résultats, et du fait qu'il leur importait peu de savoir s'ils avaient été faits par telle ou telle méthode.

Le troisième point, c'est qu'en plus, ouvrant la porte à ça, ce serait donner un *quitus* à tous les investigateurs qui ne prennent déjà pas beaucoup soin de recueillir la qualité de vie de leurs patients.

M. RUSCH.- Je remercie aussi le chef de projet d'avoir précisé que quand on a de mauvaises données à l'initial, c'est difficile de faire de la sensibilité ou des études de sensibilité dessus. Je referme la parenthèse. Cela me semble effectivement un argument important parce qu'ils avaient mis en avant qu'ils avaient fait énormément d'études de sensibilité.

Hassan.

M. SERRIER.- C'est juste pour aller dans le sens de ce qui a été dit, mais concrètement, ils n'ont pas vraiment apporté de nouveaux arguments. C'est exactement la même chose. Si l'on prend un peu de recul, c'était même un peu léger dans l'argumentaire parce que le fait d'avancer l'espoir

de guérison, sauf erreur de ma part, c'était pour expliquer le fait qu'ils aient une utilité qui est supérieure à la population générale.

Mais à chaque fois, c'était : on a une utilité qui est supérieure à la population générale, on suppose que c'est ça. Oui, mais n'empêche que, dans les faits, l'utilité est supérieure, et que c'est problématique. On ne croit pas. Ils peuvent supposer tout ce qu'ils veulent. A chaque fois, il y avait beaucoup d'arguments qui étaient tout de même très : « on suppose », « on pense que » et ça ne change pas malheureusement le fond du problème.

Personnellement, je n'ai pas été convaincu non plus. Je suis aussi, même si c'est annexe, d'accord avec Claire sur le fait qu'attention aux conséquences de valider un dossier qui a des données de qualité de vie aussi faibles. Parce que ça envoie un signal : « c'est bon, ça passe, il n'y a pas besoin d'aller plus loin ». Je redis, je pense que c'était assez clair dans mon intervention, mais que pour moi, ils pouvaient tout de même faire mieux, malgré le Covid. Je suis assez d'accord sur le fait que je ne comprends toujours pas bien. Peut-être que l'utilisation de tablettes explique un peu les choses, mais encore, même les tablettes pendant le Covid... Je ne suis pas convaincu par l'argument Covid pour le remplissage. Au contraire, pour moi, ça traduit peut-être une importance moindre de la qualité de vie qui a été délaissée à ce moment-là, mais qui aurait pu être appuyée et mieux collectée. Personnellement, cela ne change pas mon avis sur le dossier.

M. RUSCH.- On nous a indiqué que sur la question de Marc, c'était dans le dossier. Avons-nous ces éléments de réponse sur la complétude des questionnaires avant la période Covid ?

Une cheffe de projet, pour la HAS.- On a cherché dans les quelques minutes qu'on a eues après la question. Ce n'est pas flagrant. Sur la bibliographie, il y a tout un document sur l'utilité. On a effectivement les deux propositions à chaque moment, mais on n'a pas vu de graphique montrant, dans le temps, le nombre de complétions qu'il y a eues. Pour l'instant, je ne peux pas vous assurer qu'il n'y a pas eu la réponse dans le dossier, parce que bien évidemment, en dix minutes, ce n'est pas possible de re-regarder tous les éléments. Mais en regardant les documents principaux qui traitaient de l'utilité, ce n'est absolument pas mis en avant.

M. BARDOU.- C'était peut-être anecdotique, mais c'était juste pour, comme (*inaudible 1.57.11*) ce n'est pas de leur faute, (*inaudible*) être certain qu'avant la pandémie, le recueil était correct, auquel cas on peut effectivement leur donner une certaine valeur au fait qu'il y ait quelque chose, des éléments qui sont indépendants de leur volonté.

Une cheffe de projet, pour la HAS.- Par contre effectivement, ça va être un argument qui était plutôt dans leur sens. S'ils avaient eu, ou si c'était en leur faveur, ils l'auraient certainement mis en avant.

M. BARDOU.- Ils n'ont peut-être juste pas pensé, c'est possible. Je suis naïf et je leur prête un peu de bonne foi.

M. RUSCH.- Mais, l'argument et la question me semblent très pertinents. Hassan.

M. SERRIER.- C'est par rapport à ce qui vient d'être dit, je suis d'accord, ça peut apporter de l'eau moulin ou pas. Mais ça ne change rien que les données n'aient pas été collectées ou n'aient pas bien été collectées après. D'ailleurs, on peut aller encore plus loin. C'est-à-dire que sans leur jeter la faute, encore une fois, si vraiment le Covid impacte, et le Covid impacte les essais, on peut même imaginer qu'un même questionnaire rempli par un même patient avant la période Covid et pendant la période Covid, cela induise un biais parce qu'on n'a pas la même appréhension de la préférence pour son état de santé, quand on est en période de Covid ou pas, le niveau d'anxiété n'est pas le même. Et même ça, si c'était avéré, c'est un biais pour les études, et malheureusement pour beaucoup d'études. C'est pour dire qu'il y a des éléments un peu exogènes qui, même si on admet que c'est le Covid qui a causé ce problème, ça n'enlève rien au fait que les données sont difficilement utilisables derrière.

M. RUSCH.- Merci. Je propose à la CEESP de se prononcer sur le fait de maintenir notre avis initial ?

Une cheffe de projet, pour la HAS.- On peut peut-être vous présenter les différentes demandes de l'industriel sur le projet d'avis en détail, avant que vous ne votiez comme ça, on refait un tour de tout ce qui a été demandé.

M. RUSCH.- Ça me paraît tout à fait juste. Merci à vous.

Une cheffe de projet, pour la HAS.- Concernant l'analyse critique qui avait été proposée en projet initial, nous vous proposons de la garder. Nous souhaitons également ajouter un élément supplémentaire pour bien préciser que concernant les analyses de sensibilité qui ont été conduites sur le score d'utilité en survie post-événements, avec une autre source de données qui est l'essai JULIET, nous sommes toujours dans un cas de figure où nous sommes dans un essai en troisième ligne de traitement, ce qui ne permet pas d'explorer finalement le choix autour duquel il y a de l'incertitude, à savoir cette hypothèse d'équivalence entre une progression en deuxième ligne et en troisième ligne de traitement.

Enfin, concernant finalement le choix de maintenir une proposition de réserve majeure sur les utilités, nous souhaitons rappeler plusieurs éléments relatifs au choix d'une réserve majeure plutôt que deux réserves importantes. C'est notamment le fait que cette réserve majeure incombe à l'ensemble des différentes sources d'incertitudes qui ont été précédemment citées

dans notre analyse critique, et qui sont entrelées les unes aux autres et découlent toutes de la même problématique initiale, qui est en partie la collecte des données de qualité de vie.

Enfin, elle est essentiellement liée au fait que les choix et les hypothèses de modélisation, qui font l'objet finalement d'une incertitude dans les méthodes, ou les sources de données qui ont été retenues ne sont pas explorées en analyse de sensibilité, en scénario. On souhaitait bien préciser ces deux points par rapport aux données des analyses de sensibilités déterministes qui ont été présentées par l'industriel lors de l'audition.

Enfin, un dernier point que l'on souhaitait vous proposer et discuter avec vous aujourd'hui par rapport à cette réserve majeure serait peut-être une proposition de modification pour être plus clair dans le libellé de réserves majeures en proposant de modifier le début du libellé de la réserve majeure, en remplaçant « les différentes sources d'incertitude » par « l'ensemble des sources d'incertitude », pour bien montrer que cette réserve majeure est liée au cumul de ces limites sur l'estimation des scores d'utilité.

L'autre cheffe de projet va faire un tour des différents points.

Une cheffe de projet, pour la HAS.- On va aller très rapidement, mais c'est pour reprendre effectivement tous les points sur lesquels l'industriel voulait revenir.

Sur le fait d'invalider le RDCR en euros par QALY, comme on vient de le voir, effectivement, on propose de ne pas accepter cette requalification de la réserve majeure en réserve importante.

Il nous demandait de supprimer dans notre analyse critique le passage où l'on mettait en avant le fait que dès le moment de la randomisation, il y avait quasiment 50 % des questionnaires qui étaient remplis. Pour nous, cette observation reste factuelle, donc on vous propose de ne pas la retirer.

Pareil, dans notre analyse critique, il y avait tout un paragraphe où on développait l'incertitude liée à l'estimation des scores d'utilité dans l'état de santé survie sans événement, et où on disait qu'ils portaient de la moyenne des scores d'utilité de l'essai TRANSFORM à l'inclusion qui est bornée par les scores d'utilité de la population générale française. Nous proposons de garder ce paragraphe qui nous semble très important dans notre analyse critique, mais de rajouter une partie à la fin de ce paragraphe qui est dans la colonne de droite soulignée, pour bien justifier qu'*in fine*, ils utilisent les valeurs de score d'utilité de la population générale française pour qu'il n'y ait pas d'incertitude liée à ce choix.

Enfin, concernant les analyses de sensibilité, il souhaitait qu'on rajoute dans les analyses de sensibilité, dans la partie analyse critique, les analyses de sensibilité déterministes, univariées. Nous souhaitons ne pas rajouter cette partie pour rester cohérents avec ce qui est fait d'habitude

dans la synthèse des arguments proposés par l'industriel, et de ne rester que sur la présentation des analyses de sensibilité en scénario, donc de refuser cet ajout.

M. RUSCH.- Puis-je juste vous demander la première diapositive, de revenir sur la deuxième, je suppose. Celle-là, vous avez rajouté analyse de sensibilité conduite avec une autre source de données. Il s'agit de données nouvelles, ou c'était dans le dossier tout ça ?

Une cheffe de projet, pour la HAS.- Non, c'était bien dans le dossier. Cela avait été présenté dans les analyses de sensibilité en scénario. Simplement, l'industriel revenait sur le point d'avoir présenté des analyses sensibilités en scénario. Effectivement, il existe des analyses de sensibilité en scénario, mais qui ne permettent pas d'explorer les choix ou les hypothèses sur lesquelles nous avons identifié de l'incertitude(*inaudible 2.05.23*).

M. RUSCH.- D'accord. Cela me paraît clair. Y a-t-il des demandes, des questions, des interventions ? Je n'en vois pas. Je vous propose de voter sur ces modifications et sur le maintien de la réserve majeure. Samantha, pouvez-vous procéder au vote ?

(Il est procédé au vote par appel nominatif.)

M^{me} FERNANDES, pour la HAS.- 20 pour.

M. RUSCH.- Merci beaucoup.