

## 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 – DUPIXENT indiqué chez les adultes en traitement de fond additionnel de la bronchopneumopathie chronique obstructive (BPCO) caractérisée par un taux élevé d'éosinophiles sanguins, non contrôlée par une association corticoïdes inhalés (CSI), bêta-2-agoniste à longue durée d'action (LABA) et antagoniste muscarinique de longue durée d'action (LAMA) ou LABA/LAMA seule si les CSI ne sont pas adaptés (SANOFI AVENTIS FRANCE) – ECO-EFFI 796**

**M<sup>me</sup> HUIART.**- On va passer au point suivant et je vais passer la parole à Hassan sur les avis économiques, comme ça je dirai moins de bêtises.

**M. SERRIER.**- On va passer à l'avis économique DUPIXENT dans la BPCO. La présentation sera faite par les deux cheffes de projet qui sont en train de s'installer. Les rapporteurs sont Antoine et Sandrine. Vous êtes connectés ou pas ? Si vous voulez prendre un café ou autre, le temps que la connexion se fasse, c'est possible. À tout à l'heure.

*Interruption de séance.*

**Une cheffe de projet, pour la HAS.**- Bonjour à tous, nous allons vous présenter le dossier DUPIXENT, dupilumab dans le traitement additionnel de la bronchopneumopathie chronique obstructive, BPCO, caractérisée par un taux élevé d'éosinophiles sanguins, non contrôlée par l'association CSI/LABA/LAMA ou par l'association LABA/LAMA seule si les CSI ne sont pas adaptés. Nous remercions nos rapporteurs sur ce dossier qui étaient Antoine Bénard et Sandrine Loubière, ainsi que notre référent méthodologique.

Nous allons vous donner quelques éléments de contexte. DUPIXENT a obtenu son AMM en juin 2024 en procédure centralisée dans la même indication que ce présent dossier. Il s'agit d'une extension d'indication et le prix revendiqué par l'industriel est de [REDACTED] euros, toutes taxes comprises par boîte de deux stylos.

Nous avons une contribution d'association de patients dans le cadre de ce dossier. Il s'agit de l'association Santé respiratoire France qui met en avant l'efficacité limitée des bronchodilatateurs et l'absence de traitement disponible chez les patients qui ne répondent pas au traitement actuel.

L'industriel revendique un SMR important et une ASMR de niveau III. La population cible est de 38 000 patients et il ne revendique pas d'incidence sur l'organisation des soins. Le chiffre d'affaires est de [REDACTED] d'euros hors taxes en deuxième année de commercialisation. La Commission de la transparence a rendu son avis en septembre dernier et a conclu à un SMR modéré et une ASMR de niveau IV et une population cible également de 38 000 patients.

Nous allons maintenant vous parler de la BPCO. Il s'agit d'une maladie pulmonaire hétérogène qui est caractérisée par des symptômes respiratoires chroniques, tels que la production d'exacerbations ou de la toux. La BPCO ne peut pas être guérie et par définition, l'obstruction bronchique qui la définit n'est pas réversible. Son évolution peut aboutir dans les stades avancés à une insuffisance respiratoire chronique.

On distingue trois niveaux d'exacerbations : l'exacerbation légère qui est traitée par un bronchodilatateur à courte durée d'action ; l'exacerbation modérée qui est traitée par un bronchodilatateur à courte durée d'action et des corticostéroïdes oraux et parfois des antibiotiques ; enfin l'exacerbation sévère qui, elle, nécessite une hospitalisation ou une visite aux urgences.

Concernant la prise en charge, la mesure de première intention est une mesure non médicamenteuse et il s'agit de la mise en place de l'arrêt du tabac et de la réhabilitation respiratoire. Ces mesures restent à privilégier à tous les stades de la maladie.

Concernant la prise en charge médicamenteuse, en première ligne de traitement, on a les bronchodilatateurs de courte durée d'action. En deuxième ligne, on va avoir l'association LABA/LAMA et en troisième ligne, la triple association LABA/LAMA et CSI. Si, malgré cette association, on a toujours des exacerbations, on peut utiliser des macrolides ou le roflumilast.

Dans le cadre de ce dossier, il est aussi important de comprendre la classification GOLD. C'est une classification qui permet de classer la sévérité de la BPCO selon quatre catégories : GOLD 1 correspond à une BPCO légère ; GOLD 2 c'est une BPCO modérée ; GOLD 3 c'est une sévère ; et GOLD 4 c'est une BPCO très sévère. Afin de mesurer le stade GOLD, il faut mesurer tout d'abord le VEMS qui correspond au volume expiratoire maximal par seconde, qui permet de mesurer la fonction respiratoire du patient. Mais on a également une prise en compte des symptômes du patient et de son historique d'exacerbation.

Concernant les données cliniques, l'évaluation du dupilumab dans la BPCO repose sur deux études : BOREAS et NOTUS, qui ont toutes deux une méthodologie similaire. Ce sont des études de phase III, de supériorité, randomisées en double aveugle, comparatives versus placebo en ajout du traitement de fond et qui ont inclus au total 1 874 patients. Ces études avaient également des critères d'inclusion qui étaient quasiment identiques. On a juste une très légère différence sur le critère de l'inclusion relatif à l'âge des patients, mais c'est à peu près tout. La période de traitement des deux études a duré 52 semaines.

Le critère de jugement principal qui est utilisé dans le cadre de la modélisation, c'est le taux annualisé d'exacerbation modérée ou sévère de la BPCO. Les critères secondaires, de manière non exhaustive, sont : la variation du VEMS pré-bronchodilatateur à la semaine 12 et la variation

du VEMS pré-bronchodilatateur à la semaine 52. En critère exploratoire qui va également être utilisé dans le cadre de ce dossier, on a la variation du VEMS post-bronchodilatateur à la semaine 2.

Concernant les résultats, on a une différence statistiquement significative du taux annualisé d'exacerbations modérées ou sévères à 52 semaines dans le groupe dupilumab par rapport au placebo, avec une réduction du risque d'exacerbations de 30 % dans l'étude BOREAS et de 34 % dans l'étude NOTUS. Ce qui est important de comprendre pour l'interprétation de ces résultats, c'est que 90 % des événements étaient des exacerbations modérées dans chacun des groupes et surtout qu'une majorité de patients, environ 60 % dans les deux groupes, n'ont eu aucune exacerbation. Cela signifie au final, qu'il y a eu un bon contrôle de la maladie avec le traitement standard.

Je laisse maintenant la parole à l'autre cheffe de projet pour l'analyse de l'efficience.

**Une cheffe de projet, pour la HAS.-** Merci. Aujourd'hui, nous vous proposons de nous attarder surtout sur les points importants du dossier et de ne pas évoquer certains points qui nous ont semblé plutôt mineurs. Bien évidemment, si vous voulez, nous pourrions revenir sur ces points au moment de la discussion.

L'objectif de l'analyse d'efficience était d'évaluer l'efficience du dupilumab dans l'indication de demande de remboursement, ce qui paraît cohérent. Les types d'analyses proposées par l'industriel sont une analyse coût-utilité accompagnée d'une analyse coût-efficacité, dans une perspective système de soins et sur un horizon temporel de 30 ans. Cela paraît cohérent avec l'histoire naturelle de la maladie, notamment sur son côté chronique. L'actualisation était de 2,5 % pour les coûts et les résultats d'efficacité du QALY. Pour la population d'analyse, nous avons la même population que celle de la demande de remboursement.

Concernant les options comparées, dans l'intervention évaluée, il y a bien évidemment le dupilumab, et dans le bras comparateur, nous allons retrouver les associations, la trithérapie ou la bithérapie si les corticoïdes n'étaient pas adaptés aux patients.

On remarque dans les comparateurs exclus, par rapport à ce que vous a présenté la cheffe de projet, trois grandes familles, notamment le roflumilast, puisqu'il a eu un SMR insuffisant en France et n'est donc pas disponible.

Concernant les macrolides, l'industriel s'appuie sur une faible utilisation de l'azithromycine en France en s'appuyant sur l'étude BREATHE, qui est sur les données DCIR et du PMSI. Effectivement, on voit une faible utilisation de l'azithromycine en France. Par contre, l'industriel ne parle pas du tout des autres macrolides qui pourraient être indiquées chez ces patients. Du

coup, on pourrait avoir une sous-estimation de l'utilisation des macrolides, leur exclusion pourrait donc avoir un impact sur les résultats.

D'autre part, l'industriel ne prend pas en compte les mesures hygiéno-diététiques, puisque ce sont des mesures qui, normalement, doivent être instaurées dès la première ligne. Pour autant, dans les études BOREAS et NOTUS, on se rend compte que 30 % des patients inclus, parce que c'était une volonté au moment de l'inclusion des patients, qui sont encore dépendants du tabac. Le fait de ne pas instaurer ces mesures hygiéno-diététiques chez ces patients ne correspond pas aux recommandations.

Pour ces deux raisons, nous proposons une réserve mineure sur les options comparées ayant l'intitulé suivant : « La sélection des comparateurs de l'analyse n'est pas exhaustive et l'exclusion des macrolides et de l'arrêt du tabac n'est pas suffisamment justifiée. »

Concernant la population simulée, l'industriel s'appuie sur les patients des essais BOREAS et NOTUS. Pour l'analyse de la représentativité, il fait une comparaison des caractéristiques de ces données agrégées à l'étude BREATHE, qui est une étude observationnelle sur les données du SNDS chez les patients atteints de BPCO non contrôlée sous triple-thérapie.

Si on regarde un peu plus les données qui sont effectivement un peu petites, on se rend compte que les patients des essais BOREAS et NOTUS ont un profil un peu moins sévère que les patients de l'étude BREATHE. Mais ceci s'explique notamment par le fait que, comme je viens de vous le dire, les patients inclus dans l'étude BREATHE sont forcément sous trithérapie, ce qui n'est pas le cas dans les études BOREAS et NOTUS. Ceci nous paraissait acceptable.

Par contre, on n'a aucune comparaison des stades BPCO des patients des essais BOREAS et NOTUS avec des données en vie réelle. On a donc une incertitude sur la représentativité de la population simulée à la population française qui, in fine, recevra ce traitement. Pour cela, on propose une réserve mineure ayant l'intitulé suivant : « La représentativité de la population simulée à la population française ne peut être assurée en raison de l'absence des données sur la répartition des stades GOLD et de l'ensemble des comorbidités. »

Nous allons passer au modèle construit par l'industriel. Avant de rentrer dans le module 1, je vais faire une petite introduction parce que l'industriel a construit un modèle avec deux modules. Un module 1, qui est un arbre de décision, et un module 2, qui est un modèle de Markov. Pour l'instant, on va s'attarder sur l'arbre de décision et l'autre cheffe de projet vous présentera après le modèle de Markov.

Cet arbre de décision a une durée de 52 semaines, un an, et cela permet d'intégrer les données directement des essais cliniques dans ce modèle. Si on regarde en détail ce modèle 1, on voit que

les patients sont répartis en stade BPCO et qu'il y a deux durées : une première durée qui dure deux semaines et qui est une phase d'amélioration, puis une phase de stabilisation qui dure 50 semaines. À la sortie de cet arbre de décision, on aura les patients répartis dans différents stades BPCO et des patients répondeurs et non répondeurs pour l'entrée dans le modèle de Markov. Sachant qu'un patient répondeur est un patient qui a une réduction d'au moins 50 % de ses exacerbations. Les patients du bras standard et les patients ne répondant pas à ce critère dans le bras interventionnel seront considérés comme non-répondeurs.

Pour ce modèle 1, le laboratoire n'a pas suffisamment, à notre sens justifié, cette structure. Bien que cela permette d'intégrer directement les données issues des essais, il implique le recours à des données non prévues dans le plan d'analyse statistique, et nous allons revenir dessus tout de suite après pour la construction des probabilités des nœuds.

Comme je vous l'ai dit, il y a à l'entrée de l'arbre de décision, deux semaines et 50 semaines avant l'entrée dans le Markov. Or, on se rend compte que pour l'ensemble des estimations de ces probabilités de ces nœuds, l'industriel s'appuie sur les données agrégées des essais BOREAS et NOTUS, ce qui est plutôt une bonne chose, puisque cela permet d'avoir plus de patients et plus de puissance dans l'estimation de ces distributions. C'est d'autant plus le cas qu'il y a tout de même une homogénéisation des populations de ces deux essais, que ce soit en termes de caractéristiques des patients ou dans le design de ces deux études.

À l'entrée du modèle, la modélisation des patients dans les stades BPCO se construit sur les données agrégées de ces deux essais. On voit que l'industriel a modélisé des patients dans les stades BPCO légères et très sévères. Or, normalement, ces deux stades ne faisaient pas partie des critères d'inclusion dans l'entrée de ces deux essais, que ce soit BOREAS ou NOTUS. C'est un point qui est non discuté par l'industriel. Néanmoins, on se rend compte qu'il y a tout de même peu de patients dans ces deux stades.

Ensuite, à l'issue de la phase d'amélioration pour chaque bras de traitement, il y a une estimation des probabilités entre les différents stades BPCO qui s'appuie sur les données des essais BOREAS et NOTUS. Ce point nous semble peu pertinent, que ce soit cliniquement ou statistiquement. C'est un point qui n'est pas du tout discuté par l'industriel.

En effet, pour la répartition des stades BPCO, l'industriel s'appuie seulement sur la variation post-bronchodilatateur de la VEMS. C'est un critère exploratoire, qui montre une grande variabilité au cours du temps. S'appuyer à deux semaines sur ce critère est tout de même cliniquement discutable. C'est un point sur lequel on a longuement discuté, notamment avec le chef de projet médico-technique, pour être en sûr.

De plus, normalement dans les stades BPCO, il y a différents points qui rentrent en jeu et pas que le calcul de la VEMS, mais il y a également les critères associés.

Finalement construire ce module 1 sur ce paramètre ne nous semble pas méthodologiquement correct, d'autant plus qu'il n'y a aucune exploration de l'incertitude et qu'on peut s'attendre à un impact important sur le RDCR, puisque cela conditionne les stades BPCO tout au long du modèle de la phase de maintenance, mais cela influence aussi la mortalité, les exacerbations et la prise en charge des patients.

Tout cela pour vous proposer une réserve importante sur ce modèle 1 ayant l'intitulé suivant : « L'estimation des probabilités associées au nœud de l'arbre de décision à l'issue de la phase d'amélioration, à deux semaines, s'appuie sur des données non statistiquement robustes et non cliniquement pertinentes. »

De plus, dans ce module 1, les exacerbations sont prises en compte en s'appuyant également sur les données agrégées des essais BOREAS et NOTUS. Ils modélisent une différence dans les exacerbations sévères entre les deux bras. Pour autant, dans les essais, il n'y a aucune démonstration statistiquement significative qui montre un effet du DUPILUMAB sur la survenue des exacerbations sévères. Aucune analyse de sensibilité n'était associée à cette incertitude, malgré une demande réalisée lors de l'échange technique.

Pour autant, on s'attend à un impact faible de ce choix, puisque quand on regarde les données agrégées des essais, il y a peu d'exacerbations sévères qui sont survenues au cours des essais. C'est pour cela que nous vous proposons une réserve mineure sur ce point, ayant l'intitulé : « L'estimation des exacerbations sévères différentes entre les deux bras n'est pas cohérente avec les données des essais BOREAS et NOTUS. »

**Une cheffe de projet, pour la HAS.** - Je vais reprendre la suite pour le module 2. Il s'agit d'un modèle de Markov qui avait pour objectif de simuler le parcours des patients au-delà des observations cliniques, donc au-delà des 52 semaines des essais. Les patients sont, tout d'abord, répartis en fonction de leur stade de sévérité de la BPCO, entre BPCO légère, modérée, sévère et très sévère. Il n'y a pas d'amélioration du stade BPCO qui soit possible, c'est-à-dire que le patient va soit rester dans le même stade BPCO dans les cycles suivants, soit il va se dégrader. C'est cliniquement pertinent, puisqu'effectivement, en pratique, il n'est pas attendu d'amélioration du stade BPCO.

À l'intérieur de ces états de stade BPCO, on va avoir des sous-états de santé sur l'occurrence ou non d'exacerbations et sur leur sévérité. Le patient n'aura soit pas d'exacerbations, soit il aura des exacerbations modérées, soit des exacerbations sévères. Il peut avoir de une à plus de trois exacerbations à chaque fois.

L'industriel n'a pas modélisé d'exacerbations légères, et cela a été justifié, puisqu'effectivement, les exacerbations légères ont peu d'impact sur les coûts et sur la qualité de vie des patients. De plus, les essais BOREAS et NOTUS n'avaient pas pour objectif de démontrer un bénéfice sur la réduction des exacerbations légères. Cela nous semble plutôt pertinent comme modèle pour pouvoir modéliser l'histoire naturelle de la maladie. De plus, il est assez précis et permet de capturer l'interconnexion entre les différents mécanismes de la BPCO.

Sur les probabilités de transition concernant les probabilités sur les états de santé relatifs au stade BPCO, pour renseigner ces états, l'industriel va utiliser une source différente entre le dupilumab et le traitement standard durant les trois premières années du module. Pour le dupilumab, l'industriel va modéliser l'efficacité observée dans les essais BOREAS et NOTUS, alors que pour le traitement standard, il va avoir recours à des données externes. Ces données sont issues de la littérature, qui sont l'étude de Fenwick et al., étude menée au Royaume-Uni chez des patients atteints de BPCO. Ensuite, au bout de trois ans, le dupilumab va rejoindre le bras de traitement standard et les données d'efficacité vont être identiques pour les deux bras, et issues de cette étude de Fenwick.

Concernant ce choix de recourir aux données de la littérature pour le traitement standard plutôt que les données de l'essai, comme pour le dupilumab, cela n'a pas été justifié par l'industriel et cela pose problème.

Premièrement, il faut comprendre que le dupilumab est en association au traitement standard. C'est totalement incohérent d'utiliser une source différente pour le dupilumab qui est en association au traitement standard et pour le traitement standard. Il était attendu que la même source soit utilisée pour les deux bras. De plus, c'est un choix qui apparaît très largement en faveur du dupilumab, puisque dans les essais, on a observé un très bon contrôle des exacerbations avec le traitement standard. Le fait d'utiliser les données de vie réelle, les données de la littérature, ces données vont être moins favorables au traitement standard plutôt que les données de l'essai.

Nous avons réalisé une analyse de sensibilité en scénario dans laquelle nous avons utilisé les données de l'essai pour les deux bras de traitement pendant trois ans. On voit que cela a un fort impact sur le RDCR, puisqu'il augmente de 65 %. C'est pour cela que nous vous proposons une réserve importante qui a pour intitulé : « L'estimation des probabilités de transition entre les états de santé relatifs à la sévérité de la BPCO, à partir de la littérature pour le bras comparateur dans le modèle de Markov, n'est pas justifiée ni cohérente avec le recours aux probabilités de transition issues des essais BOREAS et NOTUS pour le dupilumab. »

Concernant les états de santé relatifs aux exacerbations, l'industriel utilise une source identique entre le bras dupilumab et le traitement standard seul. La source de données utilisée est celle de Whittaker, qui est une base de données au Royaume-Uni. Le recours à cette source permet notamment de prendre en compte l'historique des exacerbations des patients au cycle précédent, ce qui n'était pas possible avec les données de l'essai. Cela nous semble plutôt pertinent, puisque l'occurrence à la sévérité des exacerbations est dépendante du nombre et de la sévérité des exacerbations précédentes. Toutefois, le recours à une source externe est toujours source d'incertitude. Cette hypothèse n'a pas été testée en analyse de sensibilité en scénario, malgré une demande lors de l'échange technique.

Concernant la mortalité, l'industriel modélise deux surmortalités : une surmortalité qui est liée aux exacerbations et une surmortalité qui est liée au stade BPCO. Il n'a pas pu utiliser les données des essais cliniques, puisque ces essais étaient d'une durée de 52 semaines, ce qui est trop court pour pouvoir estimer une mortalité. Il a eu recours à des données externes de la littérature. Pour la mortalité liée aux exacerbations, c'est l'étude de Whittaker qui est utilisée. Pour la mortalité liée au stade BPCO, c'est l'étude de Shavelle et al., qui est une étude menée sur une cohorte aux États-Unis.

Le recours à la source Whittaker pour l'estimation de la mortalité liée aux exacerbations nous semble acceptable, notamment car c'est un modèle qui permettait de corriger les résultats sur plusieurs facteurs, dont le stade BPCO, afin justement d'éviter un double comptage de la surmortalité.

Cependant, le recours à l'étude de Shavelle pour l'estimation de la mortalité liée au stade BPCO nous pose plusieurs problèmes. Premièrement, c'est une étude ancienne qui date de 2009 et qui a été menée aux États-Unis. Deuxièmement, le modèle utilisé dans cette étude ne permettait pas de corriger sur les exacerbations. Cela veut dire qu'on a un risque de double comptage entre la surmortalité liée aux exacerbations et la surmortalité liée au stade BPCO. De plus, l'estimation de la surmortalité qui est associée au stade BPCO 3 et 4 repose sur une proportion très faible de patients dans l'étude, environ 3 %, alors que dans les essais BOREAS et NOTUS, on avait quasiment 50 % des patients qui étaient dans ces stades de sévérité.

L'industriel nous a tout de même proposé une analyse de sensibilité en scénario qui ne prend pas en compte la surmortalité associée au stade BPCO. Celle-ci a un impact assez important sur les résultats puisqu'elle vient augmenter le RDCR de 67 %. On vous propose une réserve importante qui a pour libellé : « Le recours aux données de l'étude Shavelle pour l'estimation de la surmortalité liée au stade BPCO est source d'incertitude, notamment étant donné le risque de double comptage avec la surmortalité liée aux exacerbations. »

Sur la gestion de la dimension temporelle, je vais passer assez rapidement. La durée de simulation est de 30 ans, ce qui est cohérent avec l'horizon temporel. Les cycles sont de un an, c'est acceptable étant donné que c'est une maladie avec une progression assez lente.

Sur l'extrapolation de l'effet de traitement, l'industriel fait l'hypothèse que les données de l'essai sont maintenues pour le dupilumab pendant trois ans, ce qui ne nous semble pas acceptable. Mais nous n'avons pas mis de réserve là, on a préféré la mettre dans la partie probabilité de transition.

Sur l'extrapolation des coûts, c'est assez conservateur, puisque les coûts vont être appliqués jusqu'à l'interruption du traitement, malgré le fait que l'efficacité du dupilumab n'est appliquée que pendant trois ans.

Sur les événements intercurrents, l'industriel sélectionne les événements indésirables sur une fréquence supérieure à 2 % dans l'un des deux bras de traitement des essais, et les événements indésirables ont un impact uniquement sur les coûts. Il faut savoir que le profil de tolérance du dupilumab est tout de même assez connu, puisqu'il a énormément d'indications. Il est favorable, donc on n'a pas vraiment de différence de tolérance entre les deux bras de traitement.

Concernant les arrêts de traitement, l'industriel modélise une probabilité d'arrêt qui est constante à chaque cycle. Nous proposons une réserve mineure, puisque quand on regarde en détail les raisons pour lesquelles le traitement a été arrêté, on voit que c'est en majorité lié à l'aversion quant à l'utilisation du dupilumab. On s'attend donc à ce qu'il y ait plus d'arrêts de traitement au premier cycle plutôt que les cycles suivants. On propose une réserve mineure qui a comme intitulé : « L'hypothèse d'une probabilité d'arrêt de traitement identique à chaque cycle n'est pas justifiée et n'est pas plausible cliniquement. »

Concernant les complications cardiovasculaires, le taux de survenue est issu des essais pour le module 1 et issu de la littérature pour le module 2.

Maintenant, nous allons passer aux utilités. L'industriel utilise les données de l'essai NOTUS, du questionnaire EQ-5D-5L. Il faut savoir que dans l'essai BOREAS, ce questionnaire n'était pas disponible, il n'a pas été proposé aux patients, donc on n'a vraiment que les données de l'essai NOTUS. L'industriel a utilisé un modèle linéaire à effet mixte pour mesures répétées. Il nous a présenté trois modèles, que je vais vous présenter juste ici. Ces trois modèles permettaient une estimation des scores d'utilité assez similaire. L'industriel a choisi le modèle qui permettait le meilleur ajustement statistique, c'est-à-dire le modèle qui permettait de minimiser les critères AIC et BIC. Au final, c'est le modèle 2 qui a été retenu.

Comme vous pouvez le voir, les coefficients associés à quatre des cinq variables incluses ne sont pas statistiquement significatifs : la variable traitement, âge, genre et état de santé. De plus, elles ne sont pas forcément cliniquement justifiées, comme la variable traitement ou la variable genre.

L'industriel ne nous a pas proposé d'analyse de sensibilité en scénario, prenant en compte un autre modèle que le modèle 2. Mais comme nous avons les scores d'utilité qui sont présentés, nous avons fait tourner le modèle avec ces scores d'utilité pour voir s'il avait un impact sur le RDCR. Comme ils sont quasiment identiques, l'impact est inférieur à 1 %. C'est pour cela que nous avons penché vers une réserve mineure, avec comme libellé : « La robustesse et la plausibilité des estimations des scores d'utilité relatifs aux états de santé ne sont pas vérifiées, en raison notamment d'une absence de significativité des coefficients associés à quatre des cinq variables de la régression retenue. Toutefois, l'impact sur les résultats est en partie exploré à travers des analyses de sensibilité en scénario réalisées par le service et est attendu faible. »

Il faut savoir que sur le niveau de la réserve, entre réserve mineure et réserve importante, nous avons eu beaucoup de discussions, que ce soit avec les rapporteurs ou avec notre référent méthodologique. En effet, méthodologiquement, ce n'est pas correct d'inclure des variables qui ne sont pas statistiquement significatives et qui, en plus, ne sont pas forcément cliniquement justifiées. Mais comme l'impact est faible sur les résultats par rapport à la doctrine de la CEESP, nous avons plutôt penché vers une réserve mineure, mais cela pourra être discuté.

Sur les désutilités, nous n'avons pas de point particulier. L'industriel a recouru aux données de la littérature pour pouvoir estimer les désutilités associées aux exacerbations et aux complications cardiovasculaires. L'industriel a mené une étude de comparabilité de ces deux études par rapport à la population simulée.

**Une cheffe de projet, pour la HAS.** - Je reprends la parole. Pour l'estimation des coûts, on vous propose de passer assez rapidement, parce qu'il y avait effectivement quelques points à discuter dans l'analyse critique, mais qui n'ont pas nécessité de réserve. Si cela vous va, on passe sur ce point, sauf si vous voulez revenir dessus au moment de la discussion.

Concernant la validation interne, l'industriel a comparé l'espérance de vie des patients modélisés aux données de la population générale de l'INSEE et également les taux d'exacerbation annualisés, modérés ou sévères, du bras de traitement standard et la répartition des patients selon le stade BPCO avec les données agrégées des essais BOREAS et NOTUS.

Pour la validation externe, différents paramètres ont été modélisés, qui ont été comparés avec des sources externes. Sur cette validation externe, on a trouvé que l'exercice n'était pas complet puisqu'il n'y avait pas suffisamment à notre sens de discussion sur cette validation externe par l'industriel, et notamment sur l'impact que pourraient avoir les différences des paramètres

modélisés par rapport aux données des sources externes. On propose une réserve mineure ayant le libellé suivant : « L'impact attendu des différences observées entre les données modélisées et les données issues des sources externes est insuffisamment discuté par l'industriel. »

Pour la validation croisée, l'industriel a comparé les choix et hypothèses faites dans ce dossier avec la méthode et les résultats modélisés dans l'avis NICE relatif au roflumilast.

Pour l'exploration d'incertitude, une des grandes limites de ce dossier est l'absence d'analyse de sensibilité en scénario, malgré des demandes que nous avons faites lors de l'échange technique. Suite à différentes discussions que nous avons eues avec nos rapporteurs, nous avons pris le chemin de faire une réserve importante spécifique au manque d'analyse de sensibilité en scénario, puisque, comme on l'a vu sur certains points, nous avons une exploration assez limitée. Beaucoup de données externes sont modélisées et ont des populations qui peuvent être hétérogènes. La réserve importante aurait le libellé suivant : « Le nombre d'analyses de sensibilité en scénario proposé par l'industriel est insuffisant et les discussions associées aux analyses réalisées sont absentes. L'exploration de l'incertitude générée par les choix de modélisation est ainsi limitée. »

Pour éviter cette histoire de réserve sur les mêmes choses, et par rapport à la réserve que nous avons vue tout à l'heure où je vous ai dit que ça ne prenait pas en compte l'absence d'analyse de scénario, nous avons fait attention dans le libellé des réserves de ne pointer l'absence d'analyse en scénario que dans cette réserve importante.

Pour les analyses de sensibilité univariée, deux problématiques ont été soulevées. Le fait que le choix des bornes des analyses de sensibilité univariée déterministes a été fixé arbitrairement et n'a pas été discuté par l'industriel. Pour le choix des distributions pour les analyses probabilistes, on aurait aussi aimé davantage de discussion. C'est pour cela que nous vous proposons une réserve mineure sur les analyses déterministes et probabilistes ayant un libellé suivant : « Les analyses de sensibilité réalisées sur les paramètres liés à l'efficacité des interventions, analyses déterministes et analyses probabilistes, sont insuffisamment commentées. »

Pour vous présenter quelques résultats de cette analyse de l'efficience, nous voyons sur un horizon temporel de 30 ans que l'ajout de dupilumab en plus du traitement standard, en comparaison au traitement standard seul, est associé à un surcoût actualisé de 65 000 euros, un gain de 0,43 QALY et un RDCR de 154 000 euros par QALY.

Pour les analyses de sensibilité, je ne reviendrai pas forcément sur toutes. Dans les paramètres qui ont un fort impact sur le RDCR, quand il est seulement modélisé une surmortalité liée uniquement aux exacerbations, il a un impact de plus de 67% sur le RDCR. On voit que la deuxième analyse de sensibilité en scénario qui a le plus d'impact sur les résultats, c'est l'analyse de

sensibilité faite par nous, le SEM, sur une estimation de l'efficacité à partir des données et des essais dans les deux bras de traitement pendant trois ans. Cela a un impact, comme vous l'a dit l'autre cheffe de projet, de plus de 65% sur le RDCR.

Pour les analyses de sensibilité probabilistes, on voit que la probabilité de 80 % pour dupilumab d'être coût-efficace est atteinte pour une disposition à payer de 190 000 euros par QALY.

Nous allons passer à l'analyse de l'impact budgétaire assez rapidement, parce qu'on s'est rendu compte qu'à part les problématiques qui sont liées à l'analyse d'efficience et qui ont une répercussion sur l'impact budgétaire, on n'avait pas d'autres problématiques spécifiques de cet impact budgétaire.

L'objectif de cet impact budgétaire était d'évaluer les conséquences financières annuelles liées à l'introduction dans le panier des soins remboursables de dupilumab dans l'indication de demandes de remboursement. La perspective était celle de l'assurance maladie obligatoire sur un horizon temporel de trois ans, avec comme population d'analyse la population d'indication de demandes de remboursement. Les scénarios comparés comprenaient : un scénario sans dupilumab ; avec seulement la trithérapie ou la bithérapie si les corticoïdes n'étaient pas adaptés ; et un scénario avec dupilumab où, en plus de ces traitements standards seuls, une association avec dupilumab était attendue.

Pour le modèle, il s'agit d'une approche de cohortes fermées. Pour l'estimation de cette population cible, l'industriel s'appuie sur différents avis CT et sur la littérature. On voit qu'on a une estimation à l'année 1 de 38 000 patients, auxquels s'ajoute le taux de croissance de la population générale en France, dans l'année 2 et l'année 3. Cette population cible estimée est cohérente avec l'avis CT, puisqu'il s'agit exactement des mêmes chiffres. L'absence de modélisation de la mortalité est acceptable au regard de l'horizon temporel retenu.

Les parts de marché estimées par l'industriel reposent sur les prévisions commerciales et sa connaissance du marché. Dans le scénario sans dupilumab, 100 % des parts de marché sont portées par le traitement standard, ce qui nous paraît acceptable étant donné l'histoire de la pathologie. Dans le scénario avec dupilumab, on voit [REDACTED]

[REDACTED]  
[REDACTED]  
On a trouvé que les analyses de sensibilité, explorant l'incertitude liée aux hypothèses de ces parts de marché, étaient insuffisamment discutées par l'industriel. Nous avons donc décidé de mener en interne des analyses de sensibilité en scénario sur ces parts de marché, et on voit que l'impact sur le RDCR est proportionnel à la variation des parts de marché.

Comme pour l'analyse d'efficience, on a souhaité pointer plutôt l'absence d'analyse de sensibilité dans une réserve propre, parce que ce n'est pas que sur ce point où il nous a manqué des analyses de sensibilité en scénario.

Pour les données intégrées, je vous propose de passer très rapidement, parce que comme je vous l'ai dit en introduction de cet impact budgétaire, l'analyse critique est transposable à celle de l'analyse d'efficience, sauf pour certains points. C'est le cas pour la mortalité, puisqu'ici, il n'y a pas de mortalité modélisée. Il y a deux réserves importantes et une réserve mineure sur les données intégrées dans l'impact budgétaire au regard de ce qui est fait dans l'analyse d'efficience.

Pour les résultats de l'AIB, concernant la population rejointe, dans le scénario avec dupilumab, le nombre de patients traités en cumulé sur trois ans est de [REDACTED] avec une majorité de patients recevant seulement la stratégie standard puisqu'ils sont quasiment [REDACTED] patients. Pour les résultats de l'AIB, en cumul sur trois ans, on remarque un total de 416 millions d'euros.

Concernant les analyses de sensibilité, on vous propose de mettre une réserve importante ayant le libellé suivant : « Les analyses de sensibilité en scénario et univariées présentées par l'industriel ne permettent pas d'explorer convenablement l'incertitude générée par les choix de modélisation. » On se rend compte que concernant les analyses de sensibilité en scénario, le minimum requis est mené par l'industriel, et concernant les analyses de sensibilité déterministes, l'ensemble des paramètres n'a pas été testé, du coup on n'a pas d'exploration de l'incertitude liée à ces paramètres.

Pour faire une synthèse des réserves dans l'analyse de l'efficience, nous avons quatre réserves importantes et sept réserves mineures, et pour l'impact budgétaire, nous avons trois réserves importantes et une réserve mineure. Je ne sais pas comment vous voulez procéder. Propose-t-on la conclusion tout de suite ?

**M. SERRIER.** - On va peut-être écouter les rapporteurs d'abord. Merci aux cheffes de projet, nous allons maintenant écouter Antoine et Sandrine dans l'ordre qu'ils veulent.

**M. BENARD.** - On a prévu que ce soit moi qui commence. Je vais porter mon propre éclairage sur les points importants du dossier. Je vais commencer par les points forts, parce que c'est vrai qu'on a tendance à focaliser sur les réserves dans les présentations, et c'est normal. Je vois deux points forts.

D'abord, les deux essais cliniques sont mis contre le placebo, ce sont des essais internationaux avec beaucoup de patients inclus, presque 2 000 au total. Ensuite le modèle de décision utilisé me paraît cohérent avec l'histoire naturelle de la maladie, que les cheffes de projet ont présentée. Il prend bien en compte à la fois les stades de sévérité de la maladie et les exacerbations, et bien

sûr les décès. Pour moi, c'est aussi un point fort du dossier. Bien sûr, le modèle, ce n'est que le squelette, après, ça dépend bien sûr des paramètres qu'on va intégrer dans le modèle pour faire l'analyse, et c'est là qu'il y a quelques points faibles, comme cela vous a été présenté.

Il y a deux hypothèses importantes qui ont été retenues dans la modélisation et qui ont un impact important sur l'estimation de l'efficacité, impact qui a été mesuré par l'industriel lui-même pour une hypothèse, et l'autre qui a été mesurée par nos chefs de projet. On verra après à quel point c'est important dans l'éclairage de la décision de la CEEPS dans ce dossier d'avoir pu explorer cette incertitude.

Le choix de ne pas recourir aux données des deux essais pour modéliser l'évolution des patients dans le groupe standard, c'est un choix important de l'industriel. La justification est que dans un essai clinique, les patients sont mieux pris en charge que dans la vie réelle, donc on a tendance à optimiser l'efficacité du traitement standard. C'est vrai, c'est mesuré en général dans les essais cliniques. Simplement, on peut aussi faire cette hypothèse pour le traitement par DUPIXENT. Sans doute, les patients dans le groupe traité par DUPIXENT dans ces deux essais étaient bien mieux pris en charge que dans la vie réelle. On se demande pourquoi l'industriel n'a pas tenu compte de la différence d'efficacité qui a été observée dans les deux essais cliniques. Ce choix va dans le sens de favoriser le traitement par DUPIXENT, puisqu'on va augmenter la différence d'efficacité entre les deux groupes comparés dans cette modélisation.

Le deuxième choix est de considérer le risque de mortalité lié à la sévérité de la BPCO d'une part et aux exacerbations d'autre part. Quand on voit à quel point c'est difficile de savoir quelles sont les causes de mortalité dans une étude clinique, différencier ce qui relève de la sévérité de la BPCO et de ce qui relève de l'exacerbation, c'est assez compliqué. Il y a un vrai risque de double comptage de la mortalité, qui va aussi dans le sens d'avantager le DUPIXENT, puisqu'il y a plus de cas de sévérité élevés dans le traitement standard, et plus d'exacerbations dans ce groupe. Ce risque de double comptage est quasiment exclusivement porté par le groupe traitement standard.

Ce sont pour moi les deux principales limites.

Il y a une troisième limite que l'on a vue qui est la simplification de la classification des patients dans la sévérité de la BPCO qui repose uniquement sur le VEMS et pas sur les exacerbations dans ce modèle. L'industriel n'a pas mesuré spécifiquement dans les deux essais cliniques les stades GOLD, il n'a mesuré que le VEMS. Si on a un conseil à donner aux industriels, c'est de préparer la modélisation dans les essais cliniques et de bien penser à mesurer tous les paramètres qui sont utiles ensuite à la modélisation dans ces essais cliniques. Au moins, dans cette modélisation, on a la chance que l'industriel ait mesuré l'utilité cette fois.

Pour conclure, je dirais que l'avantage que nous avons dans ce dossier, c'est que les cheffes de projet ont vraiment fait un travail approfondi. Elles ont repris la main sur le modèle et elles ont elles-mêmes fait les analyses de sensibilité que l'industriel n'a pas effectuées. C'est vraiment important pour éclairer notre décision, notamment le fait que l'une des deux analyses de sensibilité que l'industriel n'a pas faites a un impact important sur le RDCR avec une augmentation quasiment de 70 %. Cela vient enrichir le dossier, ce n'est pas forcément une limite, et ça nous donne des arguments supplémentaires pour prendre notre décision sur l'avis que l'on va rendre sur ce dossier. Merci beaucoup à vous deux pour le travail effectué.

**M<sup>me</sup> LOUBIERE.**- Je ne vais pas reprendre les réserves qui ont été présentées par Antoine sur lesquelles je suis d'accord. Je veux aussi appuyer ma reconnaissance du travail qui a été fait par nos deux cheffes de projet. Sur l'ensemble des réserves qui ont été faites, un gros travail effectivement en amont a été fait par ces deux chefs de projets. Si on revient sur le modèle, au départ, sans votre lecture approfondie de ce modèle, on n'aurait peut-être pas compris la subtilité de ces modules 1 et 2. En effet, ils viennent ajouter une complexité dans ce modèle. Cela permet à l'industriel de créer quelque chose à deux semaines qui peut prêter à discussion, quand on regarde bien comment les données sont utilisées et sur quels critères on obtient une amélioration de l'état des patients. Cela amène à mettre une réserve importante sur les choix méthodologiques.

Je voulais aussi revenir, même si on a mis une réserve mineure, sur la question des comparateurs. Bien sûr, c'est la première chose que l'on va regarder au début de la lecture d'un dossier sur le médicament. L'industriel nous dit que quand on prend azithromycine, c'est moins de 6 % de la population des patients qui va recevoir ce macrolide en particulier. Mais comme l'a dit l'une des cheffes de projet, au final, il y a d'autres macrolides, et peut-être qu'on serait à un pourcentage de la population qui ne serait pas de moins de 6 %, mais peut-être de 10 %, 12 %. On voit bien derrière que des comparateurs ne sont pas pris en compte. Malgré tous les efforts de nos cheffes de projet, c'est impossible de voir quel est l'impact sur le RDCR et notamment sur la frontière d'efficience. Même si, parfois, on est obligé d'accepter le chiffre annoncé par l'industriel que les macrolides sont peu utilisés en France, on est tout de même un peu sur la réserve, sur ce point.

Autre point, dont on a beaucoup parlé, sur faut-il mettre dans chaque réserve un rappel qu'en plus s'associe de l'incertitude par manque d'analyse de sensibilité ? C'était le choix initial de revenir sur chaque réserve et de compléter en disant : « et en plus, il n'y a pas l'analyse de sensibilité attendue on fait donc une importante. » Il nous a semblé important, justement, d'en faire une réserve à part entière, puisqu'il y a eu ce travail fait par les cheffes de projet.

On a déjà pointé l'absence d'analyse en scénario dans d'autres dossiers : « il n'y a peut-être pas assez d'analyse de sensibilité, on n'a pas assez d'analyse en scénario. » Mais quand on n'a pas la capacité de les faire, on ne sait pas quelle est l'importance de cette absence. Là, le travail qui a été fait fait que, justement, on voit que ça a un impact considérable sur le RDCR. C'est tout de même méthodologiquement un vrai besoin que l'industriel se prête à ce jeu de faire des analyses de sensibilité. On voit ici dans ce dossier qu'elle était faisable, parfois elles ne le sont pas, mais elles étaient faisables. Cela avait été demandé en échange technique, ce n'est pas présent. On ne va pas aller jusqu'à une réserve majeure, parce qu'on ne souhaite pas invalider ce dossier, mais c'est important, je pense, d'avoir une réserve qui porte sur l'incertitude. C'est ce que l'on a décidé en commun accord.

Enfin, cela a été rappelé lors de la conclusion dans l'avis, mais on a tout de même une évaluation qui se fait sur 30 ans. On a un gain d'efficacité de 0,42 QALY sur 30 ans, et on a un RDCR de 154 000 euros par QALY. La question, c'est : doit-on rajouter que c'est un RDCR élevé, très élevé ? Doit-on le mettre au regard, justement, du faible gain d'efficacité ? La conclusion qui va vous être présentée a été discutée. J'espère qu'elle conviendra à l'ensemble de la commission.

**M. SERRIER.**- Merci Sandrine. Je vous propose de passer à la conclusion avant la discussion pour éviter qu'elle passe tout à la fin, et surtout c'est lié. En général, quand on discute une réserve, on modifie la conclusion en conséquence, donc voyons la conclusion maintenant.

**Une cheffe de projet, pour la HAS.**- Nous allons vous la lire. Étant donné ce qui précède, la Commission d'évaluation économique et de santé publique conclut qu'au prix revendiqué de [REDACTED] toutes taxes comprises par boîte de deux, l'analyse de référence de l'efficacité du dupilumab en association au traitement standard dans l'indication aboutit à un RDCR de 150 000 euros par QALY par rapport au traitement standard seul sur un horizon temporel de 30 ans.

Ce résultat est associé à une très forte incertitude en raison de certains choix méthodologiques et hypothèses retenues par l'industriel, susceptibles d'être en faveur du produit évalué : l'estimation d'une phase d'amélioration du stade BPCO dans le module 1 du modèle qui s'appuie sur des données non robustes statistiquement et qui n'est pas cliniquement pertinente ; l'estimation des probabilités de transition entre les états de santé relatifs à la sévérité de la BPCO à partir de la littérature pour le bras comparateur dans le modèle de Markov, qui n'est pas justifiée ni cohérente avec le recours aux probabilités de transition issues des essais BOREAS et NOTUS pour le dupilumab.

Toutes choses égales par ailleurs, lorsque l'efficacité est estimée à partir des données des essais pour les deux bras de traitement pendant trois ans, le RDCR du dupilumab en association avec le

traitement standard par rapport au traitement standard seul est associé à une augmentation de près de 65 % selon une analyse réalisée par le SEM fondée sur le modèle de l'industriel. On précise bien que c'est avec le modèle de l'industriel qu'a été estimée cette analyse en scénario.

La CEESP regrette que l'incertitude générée par les choix de modélisation soit insuffisamment explorée par l'industriel, malgré des demandes d'analyse de sensibilité en scénario lors de l'échange technique.

L'impact budgétaire s'élève à 181 millions d'euros sur 3 ans pour une population rejointe estimée à [REDACTED], ce qui correspond à une augmentation de 250 % des dépenses de l'assurance maladie dans cette indication.

La CEESP souligne que le différentiel de coût total entre les stratégies comparées dans l'analyse de l'efficacité et accumulé sur 30 ans est de 65 269 euros, presque intégralement porté par le coût d'acquisition du dupilumab. Ce surcoût est à mettre en perspective avec un gain incertain de QALY de 0,42 sur 30 ans, lié à l'estimation de l'efficacité susceptible d'être en faveur du traitement évalué.

**M. SERRIER.**- Merci.

**Un intervenant.**- J'ai juste un petit commentaire sur l'impact budgétaire. Là, c'est écrit 580 millions, alors que sur la diapositive.

**Une cheffe de projet, pour la HAS.**- J'ai noté qu'il y avait une incertitude. J'ai marqué : aller vérifier, c'est entre 416 et 580.

**M. SERRIER.**- Jennifer.

**M<sup>me</sup> MARGIER.**- On nous dit que ça coûte [REDACTED] par boîte de deux, mais ça correspond à quoi en vrai, en pratique ?

**M. MERCIER.**- C'est une injection tous les quinze jours.

**M<sup>me</sup> MARGIER.**- Mais quelle est la durée de traitement en moyenne dans les faits ?

**Une cheffe de projet, pour la HAS.**- Ça s'arrête à 52 semaines.

**M<sup>me</sup> MARGIER.**- C'est un an.

**Une cheffe de projet, pour la HAS.**- Oui, pour l'instant, ils sont un an. Dans le RCP, c'est aussi ce qui est indiqué. Par contre, en raison de la chronicité de la maladie, on s'attend à ce que ça soit plus long. Si ça marche, il n'y a pas de raison d'arrêter le traitement.

**M<sup>me</sup> MARGIER.**- Mais dans les autres indications, en moyenne, je crois que c'est cinq ans l'efficacité du dupilumab. Après cinq ans, ça décline.

**M<sup>me</sup> THEBAUT.**- Ils ont fait des analyses de sensibilité, ils prennent des durées entre trois et cinq ans.

**Une cheffe de projet, pour la HAS.**- Ils ont testé deux et cinq ans de traitement.

**M<sup>me</sup> MARGIER.**- De durée de traitement ou de maintien de l'effet de traitement ?

**Une cheffe de projet, pour la HAS.**- De maintien de l'effet de traitement.

**Une cheffe de projet, pour la HAS.**- Pour le traitement, ils modélisent un arrêt de traitement annuel. On se rend compte qu'au bout de cinq ans, il y a moins de 50 % des patients qui continuent le dupilumab. Mais c'est un pourcentage : 14 % annuel d'arrêt de traitement.

**M<sup>me</sup> MARGIER.**- Pour eux, ils arrêtent l'efficacité.

**Une cheffe de projet, pour la HAS.**- Ils arrêtent l'efficacité à trois ans. Par contre, les coûts sont jusqu'à l'arrêt du traitement.

**M<sup>me</sup> TUBEUF.**- J'ai un petit commentaire sur la rédaction. Je suis juste gênée par ce « susceptibles d'être en faveur du produit évalué ». Je trouve que ça porte un jugement, que c'est biaisé, que c'est intentionnel. Ça ne rajoute rien au texte. J'aurais tendance à retirer cette partie. C'est quand tu écris : « Ce résultat est associé à une très forte incertitude liée aux choix méthodologiques et hypothèses retenus par l'industriel », de fait.

**M<sup>me</sup> MARGIER.**- Ça dépend parce que là, on sous-entend qu'on connaît le sens de la variation. Justement, ça fait partie d'une autre de nos questions qui rejoint le point que tu soulèves. C'était quoi les relations à l'échange technique ? Parce que d'habitude, quand on leur demande de faire quelque chose, ils le font quand c'est facile de le faire. Je ne dis pas que ça a été facile, mais vous l'avez fait, donc c'était faisable. Comment ont-ils présenté le fait de ne pas le faire ?

**Une cheffe de projet, pour la HAS.**- Je vais peut-être donner deux ou trois éléments de contexte sur ces essais cliniques. On en a beaucoup discuté, notamment avec le chef de projet médico-technique, parce que justement, on ne comprenait pas très bien pourquoi ils avaient fait ces hypothèses, alors que, mine de rien, au moins de faire une analyse de sensibilité, ou des analyses qui exploraient l'incertitude était possible.

Ce qu'on voit dans les deux essais, c'est que malheureusement ou heureusement, on ne sait pas, il y a eu une efficacité du traitement standard qui a été assez forte. L'industriel s'est trouvé embêté parce que la prise en charge a été optimale dans les deux bras. Le dupilumab, a priori, c'est efficace, la CT a aussi rendu son avis. Mais on pense que l'industriel voulait mettre plus en avant l'efficacité du dupilumab, et ça aurait été embêtant pour lui, de peut-être présenter une analyse de sensibilité.

Pour l'échange technique, comme vous l'avez remarqué, il y a deux modules. Pour notre première lecture et notre première analyse avant l'échange technique, on n'avait pas l'ensemble des éléments pour arriver à la finesse de l'analyse critique. Il y a des questions qu'on aurait pu poser s'il y avait eu un deuxième round d'échange technique, mais dans le process, il n'y en a qu'un seul. L'échange technique nous a permis aussi d'avoir cette finesse, notamment sur les probabilités de transition, sur l'effet traitement et tout ça.

**Une cheffe de projet, pour la HAS.-** Pour rebondir, quand on leur a demandé à l'échange technique de nous fournir une analyse de scénario qui teste ça, ils n'ont juste pas répondu.

**M. SERRIER.-** C'est une forme de réponse.

**Une cheffe de projet, pour la HAS.-** Au début, il y avait une autre question ensuite, on leur a demandé : « Pouvez-vous nous fournir une analyse de sensibilité ? ». Ils n'ont pas dit : « On ne peut pas le faire. » Ils n'ont pas rebondi dessus. On ne nous a rien fourni, on ne nous a pas expliqué pourquoi on ne nous les fournissait pas.

**M<sup>me</sup> LOUBIERE.-** Après, quand je regarde l'échange technique, 40 questions ont été posées sur l'efficience. On a un retour de réponse sur 105 pages, des réponses ont donc été apportées, mais effectivement sur des analyses qui étaient attendues... C'est pour ça que cela va dans le sens de répondre aussi à la question qui est posée de savoir que ça serait bien d'enlever une phrase qui dit que les absences analyses qui sont en plus dans le sens du traitement, tout laisse à penser qu'ils n'ont pas fait certaines analyses pour ne pas...

**M. SERRIER.-** Surtout que ce n'est pas un industriel qui n'a pas l'habitude de ce genre de choses. Si l'échange technique est aussi long, c'est que justement le dossier n'était pas suffisamment de qualité au premier dépôt.

**M. BENARD.-** En plus, on a réussi à documenter que les choix forts de modélisation étaient vraiment dans le sens de favoriser le traitement évalué.

**M<sup>me</sup> LOUBIERE.-** Le sens de variation, on sait que c'était plutôt favorable au produit.

**M. SERRIER.-** Il y avait une question.

**M. MERCIER.-** Je voulais juste apporter quelques remarques, non pas d'un pédiatre, parce que je suis pédiatre, mais un regard médical. La majorité des étiologies des BPCO, c'est le tabagisme, le tabagisme ancien. De fait, dans les critères d'inclusion de ces deux essais, BOREAS et NOTUS, il fallait au moins avoir une consommation tabagique de dix paquets-années, ce qui n'est pas considérable. Mais dans l'inclusion, 30 % des patients qui sont inclus sont encore fumeurs actifs. On se pose quelques questions sur l'ajout du dupilumab au prix où cela coûte par rapport au prix du tabac.

Le deuxième point, c'est qu'on a vu un certain nombre d'exacerbations. On s'aperçoit comme d'habitude, lorsqu'on est inclus dans un essai, en général, les choses vont mieux sur 52 semaines, c'est-à-dire sur un an. Le fait d'être inclus dans un essai fait que les patients sont mieux suivis peut-être ou sont plus attentifs.

Le troisième point, c'est qu'un certain nombre d'autres facteurs n'ont pas été pris en compte, et en particulier toute cette polémique sur l'azithromycine. L'azithromycine, c'est un antibiotique qui a été qualifié comme protégeant le poumon. On sait maintenant que l'azithromycine s'accumule dans les cellules et qu'il faut une durée de trois jours de traitement parce qu'il y a une accumulation dans les cellules de toute façon considérable, en particulier des cellules pulmonaires.

Là, il n'y a aucun détail sur les précautions contre les maladies virales qui sont les meilleures déstabilisatrices des états respiratoires un peu limités. Par exemple, nous avons appris, nous, pédiatres, que le virus de la bronchiolite, comme les grands-parents aiment bien leurs petits-enfants, ils embrassent leurs petits-enfants et c'est comme ça qu'ils attrapent un VRS forme adulte, qui peut être aussi sévère que les gripes sévères. Vous avez entendu parler de la grippe en ce moment. Actuellement, 25 % de la population, notamment à risque, qui s'est faite vacciner contre la grippe, alors que c'est tout de même la meilleure façon de prévenir les infections virales lorsqu'on a les vaccins.

Enfin, la majorité de ces patients ont été intégrés à un âge moyen de 65 ans. Je suis absolument étonné de savoir qu'on fait une projection sur 30 ans, ou ce qu'on espère d'eux, c'est qu'ils vont vivre longtemps.

Il y a tout de même un certain nombre de données insolites dans ce dossier, comme ça, à le voir de l'extérieur. En particulier, le tabagisme, c'est tout de même un élément très important. Le deuxième point, c'est le suivi des patients, puisqu'ils ont très peu de différences dans les exacerbations sévères.

Dernier point dans le papier qui a été publié dans le New England de 2024, le dupilumab n'est efficace que si véritablement on a un taux d'éosinophiles circulant qui est un marqueur un peu à la loutche d'inflammation dans lequel effectivement un anti-interleukine 14 ou 13 va être efficace. Si vous n'avez pas eu vraiment de réponse de Sanofi et de Regeneron, qui sont les deux laboratoires qui sont derrière, c'est qu'ils sont gênés aux entournures.

Voilà les petites remarques, je voulais faire comme ça en passant.

**M<sup>me</sup> THEBAUT.**- Merci beaucoup. Je voulais rebondir sur cette question du tabagisme. Je voulais savoir si vous aviez des résultats d'une analyse en sous-groupe dans l'essai sur l'efficacité sur la

population de gens qui continuent de fumer. Il me semble que cela aurait été une analyse de l'efficacité dans cette sous-population, et un élément important. Parce qu'il est possible que si l'efficacité est moins bonne dans cette sous-population, le ratio différentiel coût-résultat soit beaucoup plus élevé pour ces patients. Je crois que l'estimation de RDCR dans ces analyses en sous-groupe est possible par le guide de la HAS. Cela permet d'illustrer une vraie question d'efficacité ou de justice, qui est l'allocation de ressources pour des populations qui ne vont pas mettre en œuvre les mesures de prévention préalables. C'est une vraie question.

Après, il y a la question du contexte global. Vous l'avez dit, mais rapidement, DUPIXENT c'est un médicament qui est déjà sur le marché, dans d'autres indications, sur la dermatite atopique. En gros, on se rend compte qu'il fonctionne sur plein de trucs différents. À chaque fois, on va nous proposer des nouvelles indications. Attention, parce que dans la formulation, par exemple, dans l'impact budgétaire, vous dites que l'introduction dans le panier de soins va générer telle dépense. Il est déjà dans le panier de soins. La question, c'est à qui on va l'administrer. Le périmètre est amené à s'élargir, parce que là, cela marche sur la BPCO. On va se rendre compte que cela a peut-être un effet sur tel ou tel truc. La question, c'est vraiment : est-on prêt à payer [REDACTED] l'injection tous les 15 jours pour des preuves dont vous avez démontré qu'elles sont relativement fragiles ? En l'occurrence, il y a cette question tout de même des modes de vie. Il me semble qu'une analyse dans cette sous-population aurait été très intéressante. Je ne sais pas ce que vous en pensez et je voudrais ouvrir la discussion là-dessus.

Le dernier point, c'est que je trouve ça super que vous ayez fait l'évaluation de l'analyse de sensibilité qui n'a pas été faite. Personnellement, vous ne l'auriez pas faite, il aurait fallu mettre une réserve majeure. La raison pour laquelle je n'insisterai pas là-dessus, c'est parce que vous l'avez faite. Par contre, je trouve ça dommage que dans la conclusion, vous mettiez juste augmentation de 65 %, et que vous ne mettiez pas la valeur du RDCR parce que je trouve que ça perd en portée tout de même. Je serais très favorable au fait que vous mettiez la valeur à côté de 65 %, parce que ça montre tout l'intérêt.

**M. SERRIER.** Je vais parler généralement du service, mais vous parlez souvent en pourcentage. Dans les augmentations, je fais souvent la remarque : un pourcentage en fonction de la valeur, ça ne veut pas dire grand-chose.

**Mme THEBAUT.** - Là, on se rend compte qu'avec les 65 %, je ne sais pas à combien c'est. Quand vous nous amenez 168 000 euros par QALY, ce sont des ordres relativement raisonnables. Avec les 65 %, on commence à se rendre compte que l'injection à [REDACTED] tous les 15 jours, dans cette indication, pose une vraie question d'allocation de ressources, étant donné des réserves

que vous exprimez sur l'efficacité, etc. Il y a toujours aussi cette question de la durée de traitement et le maintien de l'effet. Il se passe des choses en termes de coût et d'efficacité.

**Une cheffe de projet, pour la HAS.-** Pour commencer sur ce dernier point, parce que la réponse est finalement la plus courte, on rajoutera le montant dans la conclusion, sauf si des personnes ne sont pas d'accord. On arrive à un RDCR de 254 000 euros, à la base, c'est 154 000 euros.

**M<sup>me</sup> THEBAUT.-** C'est raisonnable encore.

**Une cheffe de projet, pour la HAS.-** Pour rebondir sur l'efficacité en sous-groupe, on avait demandé à l'industriel, lors de l'échange technique, de nous fournir les Forest Plots que l'on vous a projetés ici. On peut voir qu'il n'y a pas de différence entre les patients qui sont toujours fumeurs et les anciens fumeurs. L'efficacité n'est pas différente.

**M<sup>me</sup> THEBAUT.-** C'est bien. On aurait un RDCR à peu près identique.

**Une cheffe de projet, pour la HAS.-** Sachant qu'ils nous ont fourni également une analyse complémentaire où ils ont intégré les coûts du sevrage tabagique.

**Une intervenante.-** Il y a une différence là tout de même, on est d'accord ? Il y en a un qui est bien à zéro et l'autre...

**Une cheffe de projet, pour la HAS.-** Ce n'est pas significatif. Il y avait aussi NOTUS et je crois que dans NOTUS, il y avait vraiment aucune différence.

**M. BENARD.-** Effectivement, il y a une analyse en sous-groupe dans les anciens fumeurs qui a inclus moins de patients, donc l'intervalle de confiance est plus large. L'estimation ponctuelle du risque relatif ça paraît beaucoup.

**M<sup>me</sup> THEBAUT.-** C'est bien d'avoir ça.

**Une cheffe de projet, pour la HAS.-** Si je peux juste rebondir sur la remarque sur les infections. Nous avons interrogé l'industriel au moment de l'échange technique, notamment pour les comparaisons des infections et la prise en charge en prévention, parce que cela fait partie des recommandations. Ils nous ont également fourni une analyse complémentaire avec la prise en charge de ces infections, à savoir le vaccin. C'est sûr que les vaccins étaient pris en compte dans cette analyse complémentaire. C'est quelque chose sur lequel nous avons essayé de les challenger un peu au moment de l'échange technique. Ils nous ont apporté des réponses a priori qui n'étaient pas si mal que ça.

**M. SERRIER.-** Il nous reste quelques minutes. Je voulais juste revenir sur quelques points pour qu'on avance. Je lance le sujet et je te donne la parole après, car il ne nous reste pas beaucoup de temps. Il y a trois points que je voulais souligner, deux qui ont déjà été discutés, un qui ne l'a

pas été. D'abord un point de forme : j'ai été très perturbé par le « non statistiquement et cliniquement pertinent ». Cette formulation m'a perturbé. N'est-ce pas juste « non robuste statistiquement » ? C'est un point de détail.

Après, je voulais revenir sur deux points. À la lecture de l'avis, je n'avais pas totalement saisi l'intérêt du dernier paragraphe. Je comprends mieux avec la présentation. Je n'avais pas saisi les messages que vous voulez faire passer. Ma question, c'est faut-il le modifier ou pas pour insister davantage ? Ce que je comprends, c'est que ce que vous voulez dire, c'est qu'en gros, le bénéfice est faible sur 30 ans et que c'est coûteux malgré tout. Je n'aime pas tourner autour du pot, si ce n'est pas suffisamment clair, il faut le dire.

L'autre point, qui est un vrai sujet, c'est revenir sur les analyses faites par le service, de la réserve importante qui a été posée. Je trouve ça très bien de poser une réserve importante. On ne l'a pas fait habituellement, mais globalement, il n'y a pas assez d'analyse. Mais tout de même, il y a un sujet derrière qui pose question parce que l'industriel est arrivé avec un dossier léger. Le service fait les analyses pour dégrader des réserves, au passage. Parce que normalement, si on n'a pas l'info, la réserve est plus importante. Je ne mets pas en cause. C'est intéressant, on a de l'information, on peut l'utiliser si cela n'a pas demandé trop de travail. Par contre, je me demande s'il ne faut pas encore être plus incisif dans le passage où l'on regrette de ne pas avoir plus quelque chose comme cela. Je ne m'en rappelle plus exactement.

**Une cheffe de projet, pour la HAS.** - Oui, c'est « La CEESP regrette que l'incertitude générale... », c'est ça ?

**M. SERRIER.** - Oui, je me demande s'il ne faut pas être encore plus incisif. Vous allez trouver que je vois le mal partout, mais qu'est-ce qui empêche un industriel de déposer un dossier léger sans mettre les analyses de scénarios dès le moment où elles sont faisables, que derrière on les fasse systématiquement et qu'ils aient une réserve importante, donc ils s'en fichent un peu.

**M<sup>me</sup> HUIART.** - C'est bien, ça réduit leurs coûts, non ?

**M. SERRIER.** - Oui, mais ça impacte le travail du service. Globalement, je suis d'accord avec la façon dont cela a été traité. Je me demande juste s'il ne faut pas être un peu plus incisif sur cette formulation.

**Une cheffe de projet, pour la HAS.** - Je ne suis pas sûre de reprendre les points dans l'ordre. Pour la pertinence clinique et robuste, effectivement, comme il s'agissait d'un critère exploratoire, on s'était dit à la base de ne partir que sur la pertinence, mais de rajouter robuste statistiquement. On en avait discuté notamment avec les rapporteurs. Je me souviens d'une longue discussion là-

dessus. On a rajouté dans la version finale, celle qu'on vous présente, le « non statistiquement robuste et non cliniquement pertinent. »

Pour la deuxième remarque concernant la phrase qu'on a ajoutée à la fin, ce sont de longues discussions qu'on a eues avec les rapporteurs, mais également avec le référent méthodologique. Actuellement, et c'est ce que j'ai compris de la volonté de cette commission de travailler par la suite dessus, on avait tout de même peu de points de repère pour qualifier un RDCR. C'était une volonté de mettre en avant les deux parties du RDCR. C'est pour cela qu'on a proposé comme cela cette phrase.

**M. SERRIER.**- Je te coupe juste. C'est juste que quand on lit rapidement cette phrase, quand je l'ai lue rapidement la première fois, je me suis dit vous donnez des infos, vous décomposez le ratio, c'est tout. Je me suis même dit pourquoi ne pas la mettre plus haut. Ne faut-il pas la reformuler pour que ce soit clair qu'on insiste sur le fait qu'on trouve que c'est un coût élevé pour un bénéfice faible ? Je ne dis pas de reformuler totalement. Peut-être que ce n'est que moi, mais à la première lecture, je n'ai pas saisi complètement le message qu'il y avait derrière. J'ai cru que c'était juste une décomposition du RDCR.

**M<sup>me</sup> LOUBIERE.**- On était mal à l'aise, je pense, de dire très élevé, alors qu'on a eu des RDCR de 2 millions et quelques. Là, on a un RDCR de 154 000. Dans quelle mesure on peut mettre très élevé/élevé ? C'était beaucoup notre tâtonnement. On s'était dit, en revenant sur des avis qui avaient été donnés précédemment, qu'une façon effectivement de détourner, c'était de pointer en disant : « Regardez, le bénéfice clinique est très faible. » Mais ce que tu viens de dire m'irait, on ne qualifie pas le RDCR, mais on dit : un coût très élevé pour un bénéfice très faible.

**M. SERRIER.**- Ce que je veux dire par là, c'est que, par exemple, au moment où on dit qu'on a gain un certain nombre de QALY de « seulement » 0,42, pour qu'on comprenne qu'il y a un petit jugement. C'est juste qu'on est entre les deux, on a l'impression qu'on dicte le truc. On a envie de dire que c'est coûteux pour ne pas gagner grand-chose, mais on laisse comme ça.

**M<sup>me</sup> HUIART.** C'est un peu bizarre de présenter ça comme ça. C'est-à-dire qu'on déconstruit le ratio, donc on se demande pourquoi on fait des ratios.

**M. SERRIER.**- Oui, ce n'est pas faux aussi. Mais, là, on est entre les deux.

**M<sup>me</sup> HUIART.**- Si on trouve que c'est plus informatif de présenter les coûts d'un côté et l'efficacité de l'autre.

**M<sup>me</sup> MARGIER.**- Non, mais tout cela peut être présent dans le tableau quand tu fais un résultat. Il y a des gens qui, peut-être, ont plus de facilité à le comprendre comme ça. Les deux sont intéressants.

**M<sup>me</sup> THEBAUT.**- Je ne suis pas sûre que quelqu'un trouve que 65 000 euros sur trente ans ça soit beaucoup.

**M<sup>me</sup> MARGIER.**- C'est facile à comprendre tout de même, comme information.

**M<sup>me</sup> THEBAUT.**- Pour moi, le message n'est pas très clair, en fait.

**Un chef de projet, pour la HAS.**- On est toujours parfois, comme dit Musset peut-être, une porte est soit ouverte ou fermée. Il n'y a pas l'entre-deux. Toujours l'entre-deux, c'est une source d'ambiguïté, ce qui nous pousse, en début de 2025, de se pencher à travailler rapidement, pour à un certain moment, éviter toutes ces conclusions. Le problème qui se pose ici, qu'on a discuté avec les rapporteurs et avec les cheffes de projets, c'est une forme, certainement, qui n'est pas optimale, mais qui fait indirectement penser qu'on voulait exprimer que ce n'est pas bon, mais on n'a pas les moyens de l'exprimer. Parce que tout d'abord, nous n'avons pas une règle et dans le passé, on ne l'a pas fait. L'expression, d'un point de vue français, est acceptable. Mais si on va sur : coûts plus élevés, mais par rapport à quoi ? On ne sait pas par rapport à quoi. Dans l'évaluation économique actuelle, sauf erreur de ma part, quand on dit quelque chose est plus élevé, c'est par rapport à quoi ? Quand on dit 0,42, c'est faible par rapport à quoi ? On n'a pas calculé la différence.

**M. SERRIER.**- Je te coupe, parce que je vais faire des propositions pour avancer. On a plusieurs points à voir et Fanny, que je n'ai pas oubliée. Je trouve personnellement que ce n'est pas complètement clair. Pour la rendre claire, il faudra qu'on aille sur un terrain glissant, donc soit on supprime, soit on remet factuellement le différentiel de coûts et le différentiel d'efficacité dans la conclusion, mais vraiment factuellement. C'est-à-dire juste au début où on présente le QALY, on dit le ratio de tant, et on met une phrase disant que le différentiel de coûts est de tant.

**M<sup>me</sup> HUIART.**- Oui, c'est ça.

**Un chef de projet, pour la HAS.**- Non, ce n'est pas possible.

**M. SERRIER.**- Dans ce cas, on le supprime.

**Un chef de projet, pour la HAS.**- On peut supprimer pour gagner du temps.

**M. SERRIER.**- C'est ce que je propose. Je ne sais pas si la Commission est d'accord.

**M. BENARD.**- Je ne suis pas d'accord.

**M. SERRIER.**- Je trouve que le message n'est pas clair. Personnellement, je ne comprends pas pourquoi ça tombe à la fin. Quand j'ai lu le truc, c'est factuel et ça décompose le RDCR. Pourquoi on ne le met pas au début de la conclusion ? C'est ce que je me dis, en fait.

**M<sup>me</sup> LOUBIERE.-** Après, il y a tout de même la dernière phrase à laquelle on tenait : « liés à l'estimation d'efficacité susceptible d'être en faveur du traitement évalué ». Je ne suis pas sûre qu'on l'ait mise ailleurs, donc on pensait que c'était important que ça y soit.

**Une cheffe de projet, pour la HAS.-** On l'avait aussi mise sous cette forme : « ce résultat est associé à une très forte incertitude et des choix susceptibles d'être en faveur du produit évalué. »

**M<sup>me</sup> LOUBIERE.-** D'accord, ça avait été mis.

**M. SERRIER.-** Sinon, ma deuxième proposition, si tu y tiens, Antoine, à le garder, parce que ça me pose un problème. Ce qui me perturbe, c'est qu'on parle du ratio au début et on retrouve ça à la fin et on a l'impression que c'est juste la décomposition. Ne peut-on pas la remonter juste dans la conclusion ?

**M<sup>me</sup> HUIART.-** Ça aboutit à un ratio de tant, et voilà.

**Une cheffe de projet, pour la HAS.-** Si je peux me permettre, parce qu'après, on finit, ce serait de prendre cette phrase et de la monter juste là, à la suite de ça.

**M. SERRIER.-** Oui, ça peut être ça. Ça dépend de ce qu'il y a après.

**Un chef de projet, pour la HAS.-** L'intention des rapporteurs, c'était de l'exprimer à la fin, faute de mieux pour attirer l'attention.

**M<sup>me</sup> HUIART.-** Cela n'attire pas l'attention parce que ce serait de la cancéro, on dirait : ça va.

**Une intervenante.-** La façon dont vous avez formulé la partie qui est importante, si j'avais lu sans écouter le débat, « l'estimation de l'efficacité susceptible d'être en faveur du traitement évalué », je n'aurais pas compris que c'était un choix de l'industriel qui entraînait un biais de surestimation de l'efficacité. Ce n'est pas assez explicite si vous voulez vraiment mettre le doigt là-dessus, surtout pour des gens qui doivent être amenés à lire que des avis individuels.

**M<sup>me</sup> MONMOUSSEAU.-** Je suis d'accord parce que le tiret donne plein d'informations. Il donne le différentiel explicite du coût, le différentiel explicite de QALY. Il précise que le gain de QALY est incertain et que l'estimation d'efficacité est susceptible d'être en faveur du traitement. On a quatre infos en quatre lignes. On ne sait pas finalement ce qu'on doit retenir de tout ça.

**M. SERRIER.-** J'ai du mal avec le message.

**M<sup>me</sup> LOUBIERE.-** C'est pour conclure justement en revenant sur de l'incertitude, une faible efficacité, et au final, un RDCR, on ne peut pas expliquer, c'est compliqué et finir en disant en plus tout cela en faveur du traitement. C'était un peu cette idée, me semble-t-il. Je parle avec Antoine et les cheffes de projet. On ne voulait pas s'arrêter dans la conclusion, on voulait faire une phrase qui concluait notre conclusion. Peut-être qu'effectivement, il faut la supprimer. Il s'agissait de

dire qu'on ne met pas de réserve majeure, mais on veut alerter tous ceux qui vont lire ce rapport : attention, il y a une efficacité qui est tout de même très faible sur 30 ans. C'est gagner si peu sur trop de temps. Oui, peut-être qu'on a un RDCR qui ne dépasse pas les millions, mais tout de même, c'est rattaché à un très faible gain d'efficacité et des hypothèses qui sont en faveur du traitement. On voulait appuyer là-dessus.

**M<sup>me</sup> HUIART.**- Pourquoi ne ciblez-vous pas sur l'efficacité ? Parce que sur les coûts, ce n'est pas choquant par rapport à ce qu'on voit d'habitude, 65 000 euros sur 30 ans, mais ça brouille le message. Dites que l'efficacité est faible, que le gain d'efficacité est faible et qu'en plus, c'est incertain et a priori surestimé. J'ai l'impression que cela reflète plus ce que vous voulez dire, plus que 65 000 euros sur 30 ans, c'est rien.

**M. SERRIER.**- C'est un peu l'idée, on tourne autour de cela. Comme il y a beaucoup d'infos et certaines infos comme le 65 000 qui ne vont pas dans le sens que l'on pense, cela noie un peu le message. Peut-être qu'il faut enlever la première partie : « La CEESP souligne que le ratio repose sur un différentiel de QALY de 0,42 sur 30 ans, qui est incertain et susceptible d'être en faveur du traitement évalué. » C'était le message qu'on voulait passer. Ce sera peut-être un peu plus clair. On a déjà dépassé le temps, mais on va revenir sur la formulation si on veut être un peu plus dur sur le truc.

Fanny en attendant tu avais une remarque.

**M<sup>me</sup> MONMOUSSEAU.**- Ce n'est pas une c'était plusieurs. Je n'ai pas vu le nuage de points qui représente les simulations de Monte Carlo, c'est dommage parce que pour moi, le nuage de points est hyper important pour justement visualiser le degré d'incertitude dans chacune des dimensions, à la fois sur le différentiel de coût et le différentiel de QALY.

Autre question : le service a réalisé des analyses de sensibilité. Est-ce propre à ce dossier ou c'est un changement de paradigme par rapport à avant ? Parce qu'il me semble qu'avant il y avait une certaine frilosité par rapport à ça.

**M. SERRIER.**- Je me permets de couper parce qu'on a déjà eu la discussion il n'y a pas très longtemps. Ce qui avait été dit, si je me souviens bien, c'était que quand c'était facile, on faisait. Mais je trouve l'évolution intéressante, parce qu'à mon sens, quand il en manque et que c'est facile, on le fait. Par contre, il faut qu'on sanctionne d'une réserve disant qu'il n'y a pas suffisamment d'analyses.

**Une cheffe de projet, pour la HAS.**- Si je peux compléter aussi, c'est parce que là aussi, on avait des intuitions. On n'aurait pas forcément pris du temps pour les faire si on n'avait pas une intuition sur le sens des analyses de sensibilité. Typiquement pour l'estimation des utilités, au

regard des scores d'utilité des différents modèles, on avait une intuition qu'il allait avoir peu d'impact et on trouvait cela dommage. Tout n'était pas parfait dans l'estimation de ces scores, mais on trouvait cela dommage pour une fois qu'on avait des scores et cela ne nous prenait pas forcément un temps fou. C'est aussi pour cela qu'on les a faits. Peut-être qu'on n'aurait pas eu cette information.

**M. SERRIER.**- Oui, je comprends. Après, c'est borderline de se dire que c'est une intuition.

**Une cheffe de projet, pour la HAS.**- C'est par rapport à l'information que nous a donnée l'industriel qui montrait déjà qu'il y avait potentiellement un faible impact et qu'on a vérifié avec des analyses de sensibilité. Ce n'est pas une intuition.

**M. SERRIER.**- Sur la question de refaire les analyses, on a déjà eu la discussion et il faudra peut-être y revenir une prochaine fois, pas dans le cadre de ce dossier, parce qu'à chaque fois, la limite est floue. Quand accepte-t-on de la refaire ou pas ? Mais je pense qu'il ne faut pas qu'on en discute aujourd'hui.

**Une cheffe de projet, pour la HAS.**- Je précise tout de même que dans le cadre du modèle Excel que nous a produit l'industriel, c'était assez simple de faire ces analyses.

**M. SERRIER.**- De toute façon, c'est sûr que c'est un critère. Il faut que ce soit simple.

**M<sup>me</sup> MARGIER.**- C'est ça le truc. Si on leur demande l'Excel, mais qu'on ne se permet jamais de faire des analyses si elles sont simples, ça ne sert à rien de leur demander.

**M. SERRIER.**- Oui, mais là, en l'occurrence, la particularité, c'est qu'ils n'ont rien fait du tout.

**Une cheffe de projet, pour la HAS.**- Il y a tout de même une limite quand c'est le SEM qui fait, et on ne l'a peut-être pas mis suffisamment en valeur, parce qu'on en a fait tout de même pas mal dans le cadre de ce dossier, c'est que ça reste sur le modèle de l'industriel.

**M<sup>me</sup> MARGIER.**- Vous ne changez pas la structure, c'est le paramètre.

**Une cheffe de projet, pour la HAS.**- On ne vérifie pas la structure du modèle en tant que tel. Ça apporte de l'information, mais cette information reste limitée parce que, on n'a pas la mainmise sur l'ensemble du modèle. On ne peut pas discuter ces résultats par rapport à l'industriel.

**M<sup>me</sup> MARGIER.**- Pour moi, l'intérêt de fournir le modèle, c'est justement ça, c'est de pouvoir changer les paramètres. Il est évident qu'on ne va pas refaire le modèle.

**M. SERRIER.**- On est d'accord, mais c'est juste que ça marche aussi dans l'autre sens. C'est-à-dire que si on n'a pas l'information, systématiquement on devrait le faire. Ce n'est pas forcément le cas. C'est là où, il n'y a pas un traitement et ça permet de dégrader une réserve parfois. Ce débat, il faut l'avoir en dehors de la question.

**M<sup>me</sup> LOUBIERE.**- Ce n'est pas forcément que dégradé, parce que là, typiquement, quand on voit qu'en faisant une analyse en scénario, les cheffes de projet montrent qu'on a plus de 65 %, là au contraire, ça nous amène à mettre une réserve.

**M. SERRIER.**- Oui, mais quand ça le fait monter, ça me gêne moins, alors que quand c'est dégradé, c'est vraiment à l'avantage de l'industriel à 100 %, alors qu'il n'a pas fait le travail.

Ta dernière remarque.

**M<sup>me</sup> MONMOUSSEAU.**- C'était sur les utilités. Les EQ-5D sont utilisés dans l'essai NOTUS. A-t-on une description un peu précise des questionnaires et à quel moment ils ont été remplis ? Est-ce proche d'une exacerbation ou pas ? Je me demandais s'il n'y avait pas un risque de double comptage de l'effet des exacerbations sur l'utilité.

**Une cheffe de projet, pour la HAS.**- Effectivement, et ça résonne avec l'ancien dossier CASGEVY où on avait un peu les mêmes problématiques. C'est le problème de ces maladies chroniques, où finalement, on ne peut pas trop savoir en fonction du recueil des données des scores d'utilité s'il y a déjà une estimation dedans. Là, pour le coup, le recueil était tout de même assez bien décrit. C'était à des temps fixes pour l'ensemble des patients. Dans le cadre de ces deux essais, comme il n'y avait pas beaucoup d'exacerbations, in fine, les exacerbations n'avaient pas un poids énorme dans le score d'utilité par l'EQ-5D-5L qui a été recueilli auprès des patients NOTUS. Sur ce point, dans ce dossier spécifiquement, ça paraît assez cohérent de dire qu'il n'y a pas un double comptage, ou que ce double comptage est très limité au vu des exacerbations.

**Une cheffe de projet, pour la HAS.**- On leur a posé la question lors de l'échange technique et, ils ont répondu que comme le questionnaire était donné à l'inclusion à six mois et à un an, ce n'était pas suffisant pour pouvoir avoir une utilité pour les exacerbations. D'où le recours aux données de la littérature et à l'application de décréments. Mais a priori, il n'y aurait pas de double comptage.

**M. SERRIER.**- Merci. Il va vraiment falloir conclure dans pas longtemps.

Le dernier point. Dans la phrase : « La CEESP regrette que l'incertitude générée par les choix de modélisation soit insuffisamment explorée par l'industriel malgré des demandes d'analyse de sensibilité en scénario lors de l'échange technique. » Première question : l'a-t-on durci ou pas ? Et comment ? Si vous êtes d'accord pour la durcir, comment peut-on la durcir ?

**M<sup>me</sup> HUIART.**- Je trouve qu'elle est assez claire. Tu veux dire quoi de plus.

**M<sup>me</sup> MARGIER.**- « On regrette », mais moi, je regrette que le RER ait eu du retard ce matin.

**M. SERRIER.**- On a déjà utilisé d'autres choses, quand on voulait être plus dur : « inacceptable ».

**M<sup>me</sup> MARGIER.**- Ça veut dire que ça permet d'avoir un refus à un moment donné aussi.

**M. SERRIER.**- Ce n'est pas tant un refus, mais c'est de dire que vous l'avez fait une fois, OK ça passe.

**M<sup>me</sup> HUIART.**- Mais ils ne savent pas quelle réserve.

**M. SERRIER.**- Une réserve importante. Ce n'est pas méchant.

**M<sup>me</sup> HUIART.**- C'est déjà pas mal.

**M. SERRIER.**- Qu'est-ce que ça fait ?

**Une cheffe de projet, pour la HAS.**- Dans CASGEVY, on avait fait une phrase comme ça sur ITT per protocole, je ne sais pas si vous vous en rappelez. Quand Hassan a fait la réflexion, je me suis dit qu'on pouvait s'inspirer de celle-là si on voulait potentiellement appuyer plus. C'était : « La CEESP insiste sur la nécessité », plus que « regrette ». Comme ça, ça reste cohérent avec d'autres formulations qu'on a eues.

**M. SERRIER.**- Si la CEESP est d'accord, voulez-vous renforcer la phrase ou pas ?

**M<sup>me</sup> MARGIER.**- Sinon « Insiste sur la nécessité », ça me va.

**Une cheffe de projet, pour la HAS.**- Ce serait « insiste sur la nécessité de mener ».

**M<sup>me</sup> MARGIER.**- ou « de réaliser ou d'explorer »

**M. SERRIER.**- Cela convient-il à tout le monde ? Dans ce cas, on va là-dessus.

**M<sup>me</sup> ALBERTI.**- Et regrette que cela n'ait pas été fait.

**M<sup>me</sup> THEBAUT.**- Non, ça va. On n'est pas non plus sur nos sentiments.

**M. SERRIER.**- Partons là-dessus. En plus, on est cohérent avec les formulations qu'on utilise habituellement, ça ira très bien. S'il n'y a pas d'objections. On est prêts pour le vote. Samantha, on t'écoute.

*(Il est procédé au vote par appel nominatif.)*

**M<sup>me</sup> FERNANDES, pour la HAS.**- J'ai 20 votes favorables.

**M<sup>me</sup> HUIART.**- OK. Bon appétit à tous. À tout à l'heure, on a une pause déjeuner un peu compacte. On va essayer de faire au mieux pour un ciblage de l'heure.