

Commission Évaluation Économique et Santé Publique

11 mai 2021

TRANSCRIPTION DES DÉBATS

AVERTISSEMENT

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

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

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

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

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

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

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

Avis économique ECO-EFFI 454 : KESIMPTA (NOVARTIS PHARMA SAS)

M^{me} PARIS.- Nous reprenons avec l'avis économique de KESIMPTA.

Un chef de Projet du SEESP.- Bonjour à tous. [REDACTED]

[REDACTED] Je profite de cette occasion pour remercier nos deux rapporteurs, Emmanuelle Fourneyron et Lionel Perrier, de leur aide précieuse pendant l'instruction de ce dossier.

Je fais une petite présentation du contexte. KESIMPTA est un traitement de la sclérose en plaques. C'est une maladie inflammatoire auto-immune dont la cible est la myéline, qui a une prévalence d'environ 2 millions de personnes dans le monde et à peu près 80 000 personnes en France. L'incidence est de 8 cas pour 100 000 habitants par an en France. C'est la première cause de handicap neurologique du sujet jeune. Il y a 3 femmes pour 1 homme dans le sexe-ratio, et le diagnostic est principalement fondé sur un faisceau d'arguments cliniques et paracliniques, notamment d'imagerie.

Il existe plusieurs types de scléroses en plaques. Dans le graphique de droite de cette diapositive, vous avez les caractéristiques. Il y a les SEP-RR, qui progressent par poussées, et pour lesquelles chaque poussée va entraîner une dégradation de l'état de santé et donc l'apparition du handicap, des SEP qui progressent d'emblée de façon un peu linéaire et des SEP qui commencent par progresser par poussées et qui ensuite évoluent vers une forme plus linéaire.

Le handicap de la sclérose en plaques est généralement décrit par l'échelle EDSS, qui est une échelle de pratique clinique allant de 0 à 10, 0 étant un score sans aucune manifestation de handicap et 10 correspondant au décès du patient. L'échelle n'est pas linéaire. Elle évolue par palier de 0,5 ou 1 point et elle est très largement utilisée dans la pathologie. Le produit est un anticorps monoclonal humanisé recombinant qui cible sélectivement les lymphocytes B exprimant le CD20. Il y a une montée en charge dans la posologie, et une administration sous-cutanée qui le distingue des autres produits disponibles.

L'AMM a été obtenue le 26 mars 2021 dans le traitement des patients adultes atteints de forme active de sclérose en plaques récurrente, donc les SEP-R, qui sont définies par des paramètres cliniques ou d'imagerie. La demande de remboursement correspond exactement au périmètre de l'AMM. L'industriel revendique un SMR important et une ASMR III dans la SEP-RR versus le tériflunomide, qui est le comparateur de l'essai clinique, et une ASMR V dans la SEP-R très active ou sévère, dans la stratégie globale de prise en charge.

La population cible annoncée par l'industriel est d'un peu moins de 50 000 patients, et le prix revendiqué est un prix de [REDACTED] euros TTC par dose, ce qui conduit à une dépense moyenne

par patient et par an de [REDACTED] euros en année 1, et à une dépense légèrement inférieure dans les années suivantes, puisqu'il n'y a pas la phase de montée en charge de la posologie.

Le chiffre d'affaires, toutes indications confondues, est annoncé à [REDACTED] d'euros TTC par an à partir de la deuxième année de commercialisation. Concernant l'évaluation de l'efficacité, les choix structurants sont assez classiques et n'appellent pour la plupart pas de commentaires particuliers. L'objectif correspond au périmètre de l'AMM. Nous avons une analyse coût-efficacité et coût-utilité, une perspective collective avec une analyse en perspective sociétale proposée en analyse de scénario. Vous voyez sur les diapositives, en italique, les résultats de l'analyse en scénario proposée. Je ne les commente pas, mais vous les avez à l'esprit. La plupart des résultats ne sont pas forcément représentés dans les résultats. L'horizon temporel retenu en analyse de référence est de 30 ans. Il varie, en analyse de scénario, entre 20 ans et la vie entière.

L'actualisation est de 2,5 % par an pour les coûts et les résultats. La population correspond à la population de l'indication, avec une analyse en sous-population pour les patients qui sont atteints de la forme spécifique SEP-RR, mais qui représentent l'immense majorité des patients inclus dans l'essai clinique. Je reviendrai sur les options comparées dans la diapositive suivante. La modélisation inclut des séquences de traitement, c'est-à-dire des traitements de deuxième et troisième lignes après le traitement évalué.

Concernant les comparateurs, les comparateurs inclus sont listés à gauche. Ils sont tout à fait conformes aux utilisations en pratique courante, donc ils n'appellent pas de commentaire particulier. Trois comparateurs des interférons sont inclus, uniquement en analyse en scénario. Les raisons pour lesquelles l'industriel les inclut en analyse en scénario sont, d'une part le fait que ces produits n'ont pas de données d'efficacité au-delà de 3 mois, or le critère de jugement qui est retenu pour la modélisation est à 6 mois. Ce sont pourtant des comparateurs qui ont été inclus en analyse de référence dans d'autres analyses, notamment les analyses de l'ICER et l'ERG.

Il était tout à fait possible de faire une analyse qui extrapolait les résultats à 3 mois jusqu'à 6 mois. Cela a été fait en analyse de scénario, et cela conduit à exclure l'ofatumumab, donc le produit que nous évaluons, de la frontière d'efficacité. Cela nous conduit à mettre une réserve importante sur le choix méthodologique de n'inclure ces comparateurs qu'en analyse en scénario.

Des comparateurs ont été exclus de l'analyse. Cela n'appelle pas de commentaire particulier puisque ce sont des molécules extrêmement peu utilisées en pratique. L'industriel appuie son évaluation sur deux essais cliniques similaires, les études ASCLEPIOS I et II, qui sont des essais double aveugle, double placebo, de supériorité versus le tériflunomide, qui incluent des patients en SEP-RR ou SEP-SP active avec un score EDSS à l'inclusion qui varie de 0 à 5,5 et un âge des patients qui varie de 18 à 55 ans.

La durée de l'essai n'était pas prédéfinie et dépendait du nombre d'événements observés, mais une durée maximale de 30 mois était prévue au protocole. La durée médiane observée est de 17,5 mois.

Le critère de jugement principal est le taux annuel de poussées confirmées. C'est un taux qui est reconstitué en tenant compte du temps passé dans l'étude compte tenu de la façon dont les événements sont observés, sans durée a priori de suivi. Les critères de jugement secondaires sont l'aggravation du score EDSS, donc du handicap, à 3 mois d'une part et à 6 mois d'autre part. Par ailleurs, l'industriel a recueilli des données du questionnaire EQ-5D à titre exploratoire. Un peu moins de 1 000 patients ont été inclus dans chacune des deux études ASCLEPIOS.

En termes de résultats, vous avez dans le tableau, dans la première colonne, les résultats sur le taux annuel de poussée, pour ofatumumab et son comparateur. Vous pouvez observer qu'il varie de 0,10 à 0,25 à l'intérieur de cet intervalle selon les 2 études. Il était prévu au protocole que les critères liés au handicap soient observés en poolant les résultats des deux études, et vous avez les résultats qui sont affichés, donc je vais me concentrer sur les résultats à 6 mois puisque ce sont ceux qui sont utilisés dans la modélisation à titre principal.

On a observé dans cet essai clinique un taux de progression du handicap à 6 mois de 8,1 % dans le groupe ofatumumab versus 12 % dans le groupe tériflunomide, et ce résultat était statistiquement significatif, avec un hazard ratio de 0,68 et un intervalle de confiance qui reste inférieur à la borne. Vous avez la représentation graphique de ce résultat, et j'attire votre attention sur les résultats à 24 mois, puisque c'est l'horizon temporel sur lequel le résultat était jugé, sur le nombre de patients encore à risque à cette échéance par rapport au nombre de patients inclus dans l'étude.

Concernant l'analyse critique du modèle économique, nous avons affaire à un modèle de Markov avec des états de santé principalement définis par la ligne de traitement. Les patients peuvent progresser dans les lignes de traitement soit parce qu'ils ont arrêté leur traitement initial quelle que soit la cause, soit parce qu'ils ont évolué sur une forme de SEP-SP, ce qui conduit automatiquement à un changement d'état de santé. Il y a des sous-états de santé à l'intérieur de chacune des lignes de traitement, qui représentent la progression dans le score EDSS. Cette progression est définie par défaut par des cohortes d'histoire naturelle de la maladie, et l'industriel vient appliquer un effet relatif des traitements qui va conduire à ralentir la progression des patients dans les scores EDSS.

La modélisation inclut en complément des événements intercurrents qui vont venir impacter uniquement les utilités et les coûts dans les états de santé, mais pas la progression dans le modèle, qui correspondent aux poussées d'une part et aux événements indésirables d'autre part.

La population simulée dans le modèle correspond à la population de l'essai clinique. Elle présente un certain nombre de décalages par rapport aux données qui sont disponibles sur les patients susceptibles de recevoir le traitement en pratique courante. En particulier, la population modélisée a tendance à avoir un handicap plus important que dans les études en vie réelle. Du coup, le score EDSS supérieur à 2 est plus fréquent dans les essais cliniques, et donc dans la modélisation, ce qui ne garantit pas la transposabilité des résultats en pratique courante. Ce paramètre a été partiellement testé en analyse de scénario, mais de façon partielle seulement parce que l'industriel a fait varier plusieurs paramètres simultanément. Nous vous proposons donc une réserve importante sur ce choix méthodologique.

La progression de la maladie nécessite 3 matrices pour faire progresser les patients dans le modèle, d'une part la progression à l'intérieur de la maladie SEP-RR, ensuite la progression de la maladie SEP-RR vers la maladie SEP-SP et enfin à l'intérieur de la maladie SEP-SP. L'industriel aurait pu utiliser une seule cohorte, la London Ontario, qui proposait des matrices pour les 3 types de progression. Il a fait le choix de panacher les cohortes utilisées, en utilisant aussi la cohorte British Columbia pour la progression dans la SEP-RR, ce qui correspond à un choix très favorable au produit, puisque cela conduit à réduire le RDCR de 77 % par rapport à un scénario dans lequel il n'aurait utilisé que la cohorte London Ontario. Nous vous proposons une réserve importante également sur ce critère.

La comparaison du traitement avec ses comparateurs repose sur une méta-analyse en réseau pour laquelle le critère de jugement principal retenu pour le modèle est le taux de progression du handicap à 6 mois, qui est généralement un critère secondaire des essais cliniques, puisque le critère principal de jugement des essais cliniques est le taux annuel de poussées et pas la progression dans le handicap, et ce critère n'est pas rapporté pour l'ensemble des essais cliniques.

L'industriel extrait également de cette méta-analyse en réseau 3 autres critères de jugement pour la modélisation : le taux annuel de poussées pour les événements intercurrents, le handicap à 3 mois et la tolérance pour estimer les arrêts de traitement par produit.

La durée de suivi dans les essais cliniques intégrés dans la méta-analyse en réseau varie de 48 semaines à 3 ans, et quelques essais ont été exclus de la méta-analyse en réseau. Il a exclu l'essai ADVANCE, qui portait sur le PLEGRIDY, car l'industriel a jugé que les résultats étaient aberrants au regard de l'expérience clinique telle que documentée par les experts cliniciens sollicités. Cet essai a été réintroduit en analyse de sensibilité, et conduit à une augmentation assez forte du RDCR de 82 %. L'interféron INCOMIN a été exclu, ce qui pose moins de problèmes d'un point de vue clinique, et des essais de non-infériorité de forme similaire ont également été exclus, ce qui n'appelle pas de commentaire puisque c'est tout à fait justifié.

Vous avez ici la présentation du réseau de la méta-analyse. L'ofatumumab est tout en haut dans le graphique en rouge, et vous voyez que la plupart des comparateurs sont reliés à l'ofatumumab de façon indirecte à travers le placebo. J'attire votre attention sur l'OCREVUS, qui est le comparateur revendiqué de l'industriel, dans le sens où il a un mode d'action similaire à son produit, et vous voyez que les liaisons qui permettent de le comparer à l'ofatumumab sont très nombreuses.

Concernant l'analyse critique de la méta-analyse, l'industriel a présenté une étude de l'homogénéité des essais, qui est un préalable à l'utilisation du HR pour établir une comparaison indirecte. On observe que certains critères sont source d'hétérogénéité a priori, en particulier le fait que tous les essais ne concernent pas la même forme de la maladie. La taille des essais est très hétérogène et on observe des différences dans les critères d'inclusion, en particulier en termes de durée antérieure de la maladie et de traitement préalable.

Nous avons des essais chez des patients naïfs et d'autres essais chez des patients prétraités. On observe également une hétérogénéité dans la définition du critère de jugement, puisque la variation du score EDSS, qui est nécessaire pour confirmer la progression du handicap et qui est le critère utilisé dans la modélisation, varie selon les essais pour les patients qui sont dans un score EDSS égal à 0 ou supérieur à 5 à l'entrée dans l'étude. Par ailleurs, la durée d'observation change selon les essais, puisqu'ils n'ont pas tous la même durée d'observation.

A posteriori, nous observons également des facteurs d'hétérogénéité entre les patients inclus :

- sur l'antériorité des premiers symptômes, ce qui correspond à l'hétérogénéité sur les critères d'inclusion, puisque cette antériorité varie de 2 ans à plus de 2 ans ;
- sur le nombre de lésions Gd+ ;
- sur le volume des lésions T2 ;
- sur la proportion de patients ayant eu un traitement antérieur, ce qui fait écho là encore aux critères d'inclusion dans les essais.

Par ailleurs, tous les essais inclus dans la méta-analyse ne rapportent pas les caractéristiques des patients sur ces critères pour lesquels on observe une variation importante.

D'autres critères de jugement ont été jugés, à l'inverse, homogènes par l'industriel. Le score EDSS moyen à l'inclusion variait quand même de 1,9 à 3,1 selon les études, et vous pouvez voir que les résultats sous placebo variaient, pour le taux annuel de poussées de 0,33 à 1,35, pour le handicap à 3 mois de 13 % à 48 %, et pour le handicap à 6 mois de 10 % à 35 %. L'industriel a considéré que cela ne remettait pas en cause l'homogénéité des essais.

Nous vous proposons d'avoir une analyse légèrement différente de cette méta-analyse. D'une part, nous regrettons qu'il n'y ait pas eu d'analyse quantitative de l'hétérogénéité des essais. Il existe des méthodes qui peuvent être mobilisées et qui n'ont pas du tout été présentées. Par ailleurs, nous considérons qu'au-delà d'une éventuelle analyse quantitative, il y a quand même une hétérogénéité partiellement explorée et qui est très forte. En particulier, pour revenir sur la façon dont est jugé le critère de résultat sur le handicap à 6 mois, l'industriel a retraité les données de ses propres essais en retenant les 3 définitions possibles du critère de jugement du handicap à 6 mois, ce qui fait varier le hazard ratio de son produit par rapport à téréflunomide de 0,54 à 0,68. Dans la comparaison versus placebo, ce hazard ratio varie de 0,43 à 0,54.

L'industriel a retenu en analyse de référence l'approche dans laquelle il recalculait ces résultats en reprenant la même façon de traiter les données que dans l'essai d'ocrelizumab. En faisant ce choix, il a privilégié l'homogénéité versus un traitement qui est quand même très éloigné de son propre produit dans le réseau de la méta-analyse en réseau, plutôt que de privilégier l'homogénéité générale du réseau en retenant la définition la plus fréquente dans les essais cliniques qui composent le réseau. Cette approche est extrêmement favorable au produit évalué, puisque s'il avait retenu une autre définition du critère de jugement, cela aurait conduit à l'exclusion de son produit de la frontière d'efficience.

Par ailleurs, une autre critique sur la méta-analyse est le fait que le critère de jugement qui est utilisé à titre principal dans le modèle est généralement un critère secondaire des essais cliniques. Ce qui nous embête, dans le fait d'utiliser ce critère secondaire des essais cliniques dans la méta-analyse, c'est que le résultat sur le critère principal des essais n'a pas été complètement rapporté pour l'ensemble des essais. On a généralement le taux annuel de poussées qui est rapporté, mais on n'a pas l'intervalle de confiance donc on ne sait pas si l'essai était conclusif sur ce critère de jugement principal avant d'envisager d'utiliser des critères de jugement secondaires.

Le critère retenu dans la modélisation n'est pas documenté dans les essais. C'est ce qui avait conduit à rejeter 3 comparateurs pour lesquels le critère de handicap à 6 mois n'était pas retenu. Enfin, on propose d'estimer que les résultats du placebo, qui est le point de référence du réseau de la méta-analyse, sont trop hétérogènes et parfois proches du résultat observé pour les traitements actifs dans l'essai pivot.

Vous avez à droite dans le tableau l'illustration de ce commentaire, puisqu'on observait pour le handicap à 6 mois une variation de 8 % à 12 % entre les deux produits de l'essai clinique, et on voit que dans le même temps les résultats du placebo varient de 10 % à 35 %, donc une amplitude beaucoup plus forte que l'effet relatif observé dans l'essai clinique. Vous avez un résultat très similaire sur les taux annuels de poussées, qui étaient des critères de jugement principaux des

essais cliniques. Sur la base de cette hétérogénéité des essais cliniques dans la méta-analyse, nous proposons une réserve majeure sur l'estimation de l'effet relatif des traitements.

Concernant la mesure des résultats de santé, des scores d'utilité sont estimés pour 3 dimensions de l'évaluation. Il s'agit d'une part de la qualité de vie liée au handicap, qui est l'estimation de scores d'utilité par classe EDSS, et qui souffre de certaines fragilités. En particulier, on a très peu d'observations pour estimer les scores d'utilité associés aux scores EDSS 7 et 8. Il a fallu faire une extrapolation à partir des données des scores inférieurs pour estimer l'utilité dans la classe 9.

Concernant la qualité de vie des aidants, qui est modélisée et prise en compte dans l'estimation des résultats de santé, elle a été déterminée via un questionnaire EQ-5D qui a été passé auprès de 200 aidants au Royaume-Uni. L'industriel ne discute pas du tout la transposition des résultats et des modalités de prise en charge de la pathologie par les aidants entre le Royaume-Uni et la France. Par ailleurs, on observe des résultats incohérents pour le score EDSS supérieur ou égal à 7, puisqu'on observe une augmentation de l'utilité des aidants, qui ne s'explique pas par des événements objectifs ou une aide supplémentaire extérieure, par exemple, pour la prise en charge des patients qui arrivent à ce stade de handicap.

Enfin, la qualité de vie liée aux poussées présente également des limites, puisqu'il n'y a qu'une seule étude qui a été proposée pour estimer cette désutilité, et la façon d'utiliser le décrétement d'utilité sur l'ensemble du cycle alors que la durée moyenne d'une poussée est plutôt estimée autour de 46 jours présente un choix favorable au produit évalué. Nous vous proposons donc une réserve importante sur ces 3 dimensions de qualité de vie.

Ce qui est plus problématique, c'est que compte tenu de toutes ces limites sur l'estimation des utilités et la qualité de vie, l'industriel ne propose quasiment aucune analyse de l'incertitude autour de ces estimations des valeurs d'utilité. Nous vous proposons donc une réserve majeure compte tenu de l'impact attendu de la mesure de qualité de vie et de la valorisation au travers des utilités sur le résultat d'efficience, puisque nous sommes dans une pathologie dans laquelle le traitement agit principalement sur la qualité de vie et pas du tout sur la survie.

Concernant les coûts, les coûts intégrés sont les coûts liés à la pathologie, avec des coûts directs, les restes à charge pour les patients et les coûts de la poussée. Nous vous proposons une réserve importante sur l'estimation des coûts, principalement liée à l'estimation du reste à charge parce que l'estimation du reste à charge est fondée sur un questionnaire en ligne qui avait inclus peu de patients et sur lequel la représentativité des patients par rapport à la population susceptible d'être traitée n'était pas très analysée. On observe un résultat très différent dans cette étude par rapport à d'autres études qui ont pu être présentées par l'industriel, avec parfois un reste à charge qui peut monter jusqu'à 45 % du coût global lié au handicap pour les scores EDSS les plus

bas, ce qui représente un taux beaucoup plus fort que ce qui a été observé dans d'autres études similaires. Nous vous proposons également une réserve importante sur cette estimation des coûts.

Par ailleurs, l'industriel a pris en compte des coûts liés aux événements indésirables, aux coûts des traitements, au suivi médical et aux coûts d'administration. L'industriel fait l'hypothèse que comme le produit est administré par voie sous-cutanée, les patients s'auto-injecteront le produit dès la deuxième injection. Ce choix méthodologique est peu étayé. Du coup, nous vous proposons une réserve mineure parce qu'on n'en attend pas un impact très important sur le résultat.

Les analyses de sensibilité sont globalement peu informatives, notamment parce qu'on a un scénario qui teste simultanément plusieurs caractéristiques de la population simulée et les séquences de traitement, avec donc plusieurs paramètres qui varient simultanément, et on ne peut pas savoir de quel paramètre vient l'écart observé sur le résultat. Le scénario qui teste une nouvelle définition du critère de progression du handicap était présenté de façon incomplète dans le dossier. Nous n'avons pas exactement l'analyse que nous aurions souhaité avoir.

D'une façon générale, les analyses de sensibilité manquent de précision, puisque les analyses en scénario réalisées ou les résultats ne sont pas complètement détaillés. Vous avez à l'écran les exemples d'analyses qui manquaient de précision. Nous vous proposons une réserve importante sur ces analyses.

Concernant les résultats observés, l'industriel aboutit à un RDCR estimé de 77 611 euros par QALY versus AVONEX, qui est le seul autre produit sur la frontière d'efficacité avec les meilleurs soins de support. J'attire votre attention sur le fait que ce ratio passe à plus de 1 million d'euros par année de vie gagnée, ce qui correspond au point que je mentionnais plus tôt, qui est que l'impact du traitement est principalement sur la qualité de vie et pas sur la survie, ce qui explique cet écart très important entre le critère de jugement par QALY et par année de vie gagnée, ce qui renforce la nécessité d'avoir des scores d'utilité qui soient particulièrement solides ou explorés en analyse de sensibilité.

On observe que la frontière d'efficacité est rapidement modifiée en cas de variation du prix d'ofatumumab. Vous pouvez observer sur le graphique que les produits qui sont exclus de la frontière d'efficacité n'en sont pas très éloignés, et dès que l'on fait varier des paramètres, on se rend compte que la frontière d'efficacité est très instable, et les produits rentrent et sortent sur la frontière d'efficacité. L'industriel a essayé d'estimer un rapport entre le RDCR et le prix de KESIMPTA, mais il faut prendre cette estimation avec beaucoup de précautions puisque le comparateur change régulièrement. Il s'agit donc d'un RDCR estimé, mais qui n'est pas toujours en relation avec le même comparateur.

Concernant l'exploration de l'incertitude, l'industriel ne dépasse jamais la probabilité de 39,1 % d'être le produit le plus coût-efficace, en l'occurrence de maximiser le nombre de simulations dans lesquelles il apparaît le plus fréquent coût-efficace. Il atteint ce taux de 39,1 % à peu près à 150 000 euros par QALY et ne le dépasse jamais, jusqu'au seuil de 550 000 euros par QALY qui a été testé par l'industriel.

Les résultats se stabilisent relativement vite à partir de cette valeur de 150 000 euros par QALY entre les comparateurs. Les paramètres qui influencent le plus les résultats sont d'une part le HR d'ofatumumab et le HR de son comparateur, et dans une moindre mesure les valeurs d'utilité.

Concernant l'analyse d'impact budgétaire, l'industriel estime la population prévalente à 48 750 patients et une population incidente de 3 000 patients par an. L'industriel estime la population générale et applique ensuite les parts de marché estimées année par année. Il déduit de cette estimation la progression des patients sous traitement, et en particulier les switches d'un traitement à l'autre, de façon à atteindre cette part de marché cible et cette population cible. L'industriel fait l'hypothèse qu'aucun patient ne reste sans traitement et que les parts de marché sont les mêmes parmi les patients incidents et prévalents.

Il fait une troisième hypothèse qui nous paraît très importante et sur laquelle nous souhaitons attirer votre attention. C'est le fait que [REDACTED]

[REDACTED]

Nous vous proposons une réserve importante sur ce choix méthodologique qui n'est pas du tout argumenté. Le traitement des patients changeant de traitement est donc défini de façon à atteindre des populations prédéfinies par les parts de marché auparavant. Nous vous proposons également une réserve importante puisqu'il n'y a pas du tout d'analyse dynamique du switch sur le fait que le choix du traitement de deuxième ligne puisse éventuellement dépendre du traitement reçu en première ligne par le patient.

Vous avez à l'écran les résultats de l'analyse d'impact budgétaire année par année pour les 5 premières années de commercialisation. Vous pouvez observer qu'en année 5, l'ofatumumab atteint une population traitée d'environ [REDACTED] patients sur un total de 63 750 patients. L'impact budgétaire cumulé à 5 ans est d'environ [REDACTED] d'euros. Cette augmentation correspond à une variation inférieure à 1 % des dépenses par rapport au scénario sans KESIMPTA, ce qui s'explique par les hypothèses présentées plus tôt sur le fait que les prises de parts de marché se

font [REDACTED] et donc le seul impact budgétaire est la différence de coût de prise en charge entre ces deux traitements très proches.

Concernant la proposition d'avis de la CEESP, je vous présente dans un premier temps la contribution de l'association Ligue française contre la sclérose en plaques, qui a insisté beaucoup sur

- la fatigue liée à la maladie et sur la variabilité des symptômes autres que la fatigue, qui est une variabilité très forte à la fois entre les patients, mais aussi une variabilité pour un même patient, avec une fluctuation et une invisibilité des symptômes qui rend la maladie particulièrement difficile à vivre pour le patient et pour son entourage, que ce soit un entourage proche ou professionnel ;
- l'impact important de la maladie sur les proches ;
- l'importance de limiter les poussées, source de handicap ;
- la qualité de vie associée au traitement indépendamment de l'effet du traitement sur la maladie elle-même ;
- les échappements thérapeutiques fréquents observés par les patients.

Concernant l'efficience, nous vous proposons 3 réserves majeures :

- une réserve sur la définition retenue pour le critère d'efficacité CDP-6, donc le handicap à 6 mois, qui est alignée sur l'étude d'ocrelizumab, qui est à l'origine d'une très forte hétérogénéité dans la méta-analyse et qui a un impact déterminant sur l'efficience d'ofatumumab ;
- une réserve majeure sur l'absence de comparaison robuste du fait de l'absence de comparaison directe versus les comparateurs pertinents et de l'irrecevabilité de la méta-analyse en réseau proposée ;
- une réserve majeure sur l'analyse très insuffisante de l'incertitude associée aux utilités, dans un contexte où le niveau de RDCR est grandement déterminé par la qualité de vie.

Il y a également 9 réserves importantes. Nous vous proposons, en conclusion sur l'efficience, que les limites de l'évaluation proposée ne permettent pas de valider le niveau de RDCR estimé par l'industriel et empêchent de valider une estimation avec un degré de confiance suffisant. Néanmoins, l'évaluation permet de dégager des enseignements généraux sur l'efficience du traitement, avec 3 principales conclusions, au-delà du RDCR qui ne peut pas être retenu :

- le fait que l'efficacité relative des traitements est mal connue ;

- le fait que l'efficience repose principalement sur un gain en qualité de vie, et nous attirons l'attention en particulier sur le gain en qualité de vie généré par les traitements et la valorisation des préférences associées aux états de santé ;
- le fait que malgré les nombreuses limites de l'évaluation et le cumul de choix méthodologiques très favorables à l'ofatumumab, plusieurs analyses en scénario font apparaître le produit comme dominé.

On ne peut pas exclure le risque d'ofatumumab ne soit pas efficient au prix revendiqué.

Concernant l'AIB, il y a deux réserves importantes que je viens de présenter, je ne vais pas revenir dessus. Les conclusions reprennent les éléments présentés plus tôt sur le fait que l'impact budgétaire est estimé autour de [REDACTED] d'euros à 5 ans, avec une augmentation très faible, et que cet impact est très déterminé par le choix fait par l'industriel de présenter une analyse dans laquelle il fait l'hypothèse que l'ofatumumab ne vient prendre des parts de marché [REDACTED]

En conclusion générale de la CEESP, nous proposons la conclusion suivante. L'incertitude associée à l'évaluation est trop forte pour estimer le RDCR avec un degré de confiance suffisant. Le bénéfice des traitements, et donc l'efficience potentielle, repose de façon importante sur le gain en qualité de vie, et une attention particulière doit être portée à la valorisation des préférences associées à la qualité de vie. Le risque d'ofatumumab soit dominé au prix revendiqué ne peut pas être exclu. L'impact budgétaire d'ofatumumab repose sur l'hypothèse d'une prise de parts de marché [REDACTED]. L'incertitude dans l'estimation du reste à charge semble très élevée au regard des études existantes dans la littérature.

Considérant les sources d'incertitude identifiées, les résultats de l'analyse d'efficience ont besoin d'être corroborés par des données recueillies en vie réelle, visant notamment à documenter :

- l'efficacité relative des traitements ;
- l'effet à long terme du produit sur l'évolution des patients dans le handicap ;
- la valorisation des utilités des patients et des aidants associées au handicap et aux poussées ;
- le reste à charge pour les patients.

Je vous remercie de votre attention.

M^{me} PARIS.- Merci beaucoup. Les rapporteurs sont Emmanuelle Fourneyron et Lionel Perrier. Qui souhaite parler ?

M. PERRIER.- Emmanuelle, veux-tu commencer ?

Mme FOURNEYRON.- Vas-y, si tu veux.

M. PERRIER.- Bien. Je voulais remercier les chefs de projet pour le travail effectué dans le cadre de ce dossier, qui présente un certain nombre de manques et une grande fragilité. Je vais en citer quelques-unes, qui ont déjà été largement développées.

Il y a l'utilité, dont l'incertitude n'a été que partiellement explorée alors que son impact sur le ratio différentiel coût-résultat est majeur. Pour rappel, l'écart observé en années de vie gagnées entre les différentes stratégies étudiées est très faible, donc la qualité de vie doit être évaluée avec la plus grande rigueur.

Ensuite, nous notons d'autres incertitudes, comme par exemple l'effet relatif des traitements, où le niveau d'incertitude demeure important. Je citerai l'hétérogénéité des essais dans la méta-analyse, qui pose la question d'ailleurs plus générale du design des essais cliniques, et la modification, cela a été également montré, de la frontière d'efficience en fonction des hypothèses retenues.

Sur la forme, je pense que le dossier aurait à gagner en clarté. C'était un dossier un peu difficile pour vraiment entrer à l'intérieur, du fait d'un certain nombre de facteurs comme la multiplicité des sources de données ou de points insuffisamment développés dans certaines parties, même si nous comprenons bien la difficulté et la complexité de cette pathologie. Je vous remercie.

M^{me} PARIS.- Merci. Emmanuelle ?

M^{me} FOURNEYRON.- Bonjour à toutes et tous. En complément de ce qu'a présenté Lionel, je souhaite à mon tour saluer et remercier les chefs de projet pour le travail sérieux qu'ils ont fait sur un dossier qui questionne beaucoup méthodologiquement. Merci aussi pour leur disponibilité, parce qu'ils ont pris le temps d'un échange bilatéral avec moi en sachant que j'avais besoin de compléter certaines informations, donc merci beaucoup pour cela.

Que vous dire ? C'est un dossier sur la sclérose en plaques, dont la nouveauté principale me semble surtout résider dans le mode d'administration en voie sous-cutanée. Ceci étant, comme l'a très bien montré cette analyse faite par les chefs de projet, c'est un dossier qui présente de lourdes lacunes méthodologiques, et qui de ce fait nous empêche de faire notre travail, c'est-à-dire de pouvoir donner un vrai avis d'efficience sur le produit et le dossier qui nous est présenté.

De mon point de vue, la seule chose que j'ajouterais par rapport au rapport de Lionel et la présentation des chefs de projet, c'est que cela interroge quand même un peu parce que d'abord, on connaît ce laboratoire, et notamment son expertise en matière d'évaluation médicoéconomique. C'est vrai que de par ce niveau d'expertise, cela interroge un peu. Cela interroge aussi en sachant que si j'ai bien compris, il y a eu un processus d'échanges pré-dépôt,

qui aurait dû normalement éviter d'arriver justement sur ces difficultés méthodologiques dans l'avis d'efficience. Cela questionne un peu ces aspects. Dans ce contexte, je soutiens l'avis et l'ampleur des nombreuses réserves qui sont proposées dans l'avis me semble raisonnable, de même que l'importance des données complémentaires qui sont demandées dans la conclusion de l'avis. Je vous remercie.

M^{me} PARIS.- Merci beaucoup. Veux-tu répondre ?

Un Chef de Projet du SEESP.- Oui, je voulais juste compléter parce que c'est un point que je n'ai pas évoqué dans la présentation, mais ce dossier n'a pas fait l'objet d'un échange technique. Vous avez remarqué qu'il y avait beaucoup de sources d'incertitude et de réserves qui pouvaient interpeller sur le fait qu'elles auraient pu être davantage documentées. Quand nous nous sommes posé la question de l'échange technique, nous avons estimé qu'au moins une des réserves majeures, en particulier celle sur l'irrecevabilité de la méta-analyse, ne pouvait pas être levée au cours d'un échange technique. Nous avons débattu en interne sur la question de savoir s'il était opportun ou non de faire un échange technique et nous avons choisi de ne pas le faire en nous disant qu'il n'y avait pas lieu de faire un échange technique si nous savions d'emblée que nous n'arriverions pas à lever la réserve majeure sur l'effet relatif des traitements.

M^{me} PARIS.- C'est entendu, merci. Je ne sais pas s'il y a d'autres commentaires. Pour rebondir sur un commentaire que Lionel a fait, je ne sais pas si c'est exactement l'idée qu'il souhaitait exprimer mais nous pouvons quand même nous étonner qu'il n'y ait pas plus de données comparatives sur les traitements de la sclérose en plaques, puisqu'il en existe plusieurs. En tout cas, c'est un peu étonnant, mais c'est peut-être une question de calendrier. Peut-être qu'il n'y a pas d'essai avec les comparateurs les plus pertinents. Il y a quelque chose qui m'échappe un peu dans ce dossier. C'est peut-être une question de chronologie des mises sur le marché des différents comparateurs, en fait.

Un Chef de Projet du SEESP.- Il y a quand même des produits relativement anciens dans le réseau. Ils ne sont pas tous récents.

M. PERRIER.- C'était exactement mon point de vue, Valérie, c'est-à-dire privilégier les comparaisons directes.

M^{me} PARIS.- C'est clair. Catherine ?

M^{me} LE GALES.- Oui, merci beaucoup. C'est effectivement un dossier qui ne permet pas de conclure et qui a beaucoup d'insuffisances méthodologiques. Je me posais une question. Il y a deux choses. En lisant la contribution de patients, mais peut-être que j'ai un biais de lecture, j'ai été troublée par le fait qu'il n'y avait pas de référence à ce produit en particulier. Est-ce que je

me trompe ou est-ce qu'en fin de compte c'est bien le cas ? C'était ma première question, parce que c'est vrai que l'association semble exprimer vraiment un désarroi important et une demande importante de la part des malades pour des traitements qui répondent à leurs préoccupations. Pourtant, ce n'est pas le premier traitement que nous voyons. Cela interroge aussi sur les derniers traitements entrés sur le marché dans cette indication, mais peut-être que je n'ai pas bien lu et si c'est le cas, merci de rectifier mon erreur.

L'autre chose qui m'a marquée, en lisant le rapport, c'est que dans l'impact budgétaire, il s'agit très clairement d'un positionnement en concurrence [REDACTED]. J'ai bien vu qu'OCREVUS était dominé au sens de la dominance étendue, etc., mais pourrions-nous dire des choses dans l'avis d'efficience ? Nous pourrions imaginer que le fait de mettre ce produit à disposition soit aussi important pour les malades qui, pour une raison ou pour un autre, ne peuvent pas prendre d'OCREVUS ?

Il me semble que le positionnement du produit que nous examinons aujourd'hui, y compris en termes de prix, est important. Y a-t-il plus de choses que ce que vous avez dit, y compris d'ailleurs qui seraient directement utiles pour le CEPS dans la négociation du prix et que nous pourrions rajouter dans l'avis pour le positionner par rapport OCREVUS ?

J'ai compris qu'il était probablement dominé, que nous avons des réserves majeures et qu'il fallait donc être extraordinairement prudents, mais je me demandais s'il y avait des choses que nous pourrions dire par rapport à OCREVUS parce que nous pouvons imaginer, même si le produit n'est pas efficient, qu'il sera néanmoins remboursé et qu'un prix sera fixé.

Un Chef de Projet du SEESP.- Sur la contribution de patients il faudrait que j'y revienne, parce que j'avoue que je ne l'ai plus suffisamment en tête pour m'en souvenir précisément.

Concernant la comparaison à OCREVUS, c'est compliqué parce que nous voyons que sur la frontière d'efficience, ils sont très proches l'un de l'autre.

Nous serions donc tentés de dire que c'est soit l'un soit l'autre qui est sur la frontière, ce qui ferait donc sortir l'autre, et nous voyons plusieurs scénarios sur lesquels nous observons ce phénomène, donc ils sont sans doute très proches, mais cela repose sur une comparaison tellement indirecte, et que par ailleurs nous considérons comme n'étant pas du tout validée, qu'il est compliqué de s'avancer sur ce terrain. Le réseau pour relier OCREVUS à KESIMPTA est extrêmement étendu. Il passe par un placebo, et nous voyons que les placebos n'ont probablement pas été mesurés dans des populations comparables dans tous les essais. Il paraît délicat de s'avancer sur ce terrain-là.

M^{me} LE GALES.- Si je peux juste poursuivre ma question, je suis bien consciente que je ne suis pas orthodoxe dans ce que je vais dire et je ne dis pas qu'il faut la faire, mais si nous avons besoin d'une analyse complémentaire qui soit une comparaison qui se limiterait à OCREVUS dans le cadre de l'évaluation économique, pourrions-nous la faire, ou n'avons-nous aucun moyen parce que nous restons totalement dépendants de cette comparaison indirecte ? Vois-tu ma question ?

Un Chef de Projet du SEESP.- C'est-à-dire avec les données disponibles ?

M^{me} LE GALES.- Non, je me demandais si cette difficulté que nous avons à le positionner, et cette grande incertitude qui entoure le positionnement relatif de KESIMPTA par rapport à OCREVUS, est liée au fait que nous avons une comparaison indirecte qui implique beaucoup de comparateurs. Si c'est le cas, et si nous nous limitons à une comparaison à OCREVUS dans une analyse pas du tout orthodoxe, serions-nous capables de dire quelque chose en termes d'analyse économique, ou est-ce que la comparaison à OCREVUS ne pourrait se faire que de manière indirecte et donc avec les limites des données dont on dispose ? En gros, est-ce que le réseau serait le même ?

Un Chef de Projet du SEESP.- S'il fallait relier ces deux points par le réseau, il n'y a pas d'autre essai. L'exhaustivité des essais disponibles pertinents pour relier les deux produits a été présentée.

M^{me} LE GALES.- On ne peut vraiment rien faire.

Un Chef de Projet du SEESP.- Je ne crois pas.

M^{me} LE GALES.- C'était juste de la curiosité. C'est vrai que le positionnement du produit par rapport à OCREVUS, je pense, est l'un des éléments qui seront pris en compte par le CEPS.

Un Chef de Projet du SEESP.- Avec la difficulté qu'en plus, dans les deux essais disponibles, les critères de jugement ne sont pas les mêmes pour juger du critère d'intérêt entre les deux essais. Pour le coup, c'est une analyse qui a été explorée dans l'analyse en scénario. En changeant le critère de jugement, on voit que les résultats sont très instables et que le produit devient dominé. Nous savons qu'il y a une incertitude. Il n'y a pas de doute là-dessus, mais nous n'identifions pas de moyen de la réduire.

M^{me} LE GALES.- Pour pousser jusqu'au bout, je vois que nous n'avons pas de moyen de la réduire, mais aurions-nous plutôt confiance dans le fait que rien ne justifie que le prix de ce nouveau produit soit supérieur au prix [REDACTED], ou est-ce que même cela nous ne pouvons pas le dire ?

Un Chef de Projet du SEESP.- C'est une bonne question pour la doctrine de la CEESP. En tout cas, il n'y a aucun élément qui montre qu'au prix revendiqué, ofatumumab puisse être plus efficace qu'OCREVUS avec un degré de certitude suffisant. Je ne sais pas si cela vous apporte l'information.

M^{me} LE GALES.- Je pense, oui.

M^{me} PARIS.- Merci. Je ne sais pas si tu vois la remarque qui est dans le tchat.

Un Chef de Projet du SEESP.- Oui.

M^{me} PARIS.- C'est une question de formulation. « L'efficacité relative est-elle mal connue ou mésestimée ? » Ce n'est pas tout à fait équivalent.

Un Chef de Projet du SEESP.- Non, c'est vrai. Elle est sans doute mal connue parce qu'elle est mésestimée, mais je pourrai veiller à employer les termes les plus appropriés.

M^{me} PARIS.- Il y a une chose sur laquelle tu es revenue pendant la présentation, et qui m'avait un peu frappée dans le dossier. Ce sont les paiements à la charge des patients. Pourrais-tu juste réexpliquer ? J'avais retenu, de la lecture du dossier, que d'après ce que nous savions les restes à charge étaient quand même très élevés. Après, je ne sais pas si c'est du reste à charge sur le soin ou si ce sont des paiements de prestations qui ne sont pas couvertes par l'Assurance maladie, puisqu'il s'agit probablement de problèmes liés au handicap.

Un Chef de Projet du SEESP.- Oui, c'est bien le reste à charge sur la compensation du handicap qui est très fort, mais cela a un impact très fort sur le résultat, qui est une partie importante et significative du coût.

M^{me} PARIS.- Mais qui est inclus dans l'analyse, pour le coup.

Un Chef de Projet du SEESP.- Oui.

M^{me} PARIS.- C'est entendu. En tout cas, c'est vrai que c'était une donnée de ce dossier assez frappante, puisque finalement cela fait des coûts assez élevés pour les patients.

Un Chef de Projet du SEESP.- Oui, avec une part du reste à charge beaucoup plus forte dans la source retenue par l'industriel pour ce dossier que ce que l'on peut observer dans d'autres études, et il n'a pas du tout justifié le fait d'avoir utilisé cette source, dans laquelle le reste à charge est particulièrement élevé, par rapport à d'autres sources de données.

M^{me} PARIS.- Mais cela n'avantage pas particulièrement, si cela augmente le coût.

Un Chef de Projet du SEESP.- Cela peut avantager si le fait de moins progresser dans le handicap entraîne moins de coûts.

M^{me} LE GALES.- Oui, bien sûr, s'il y a moins de handicap, et donc moins de coûts.

Un Chef de Projet du SEESP.- Son produit entraîne moins de handicap.

M^{me} PARIS.- Si ce moindre handicap est montré, pour la partie que l'on peut considérer comme démontrée, c'est effectivement moins de coûts.

Un Chef de Projet du SEESP.- Oui, et cela ne joue que sur l'efficacité. Cela ne joue pas sur l'impact budgétaire, puisque ce sont des coûts qui ne sont pas inclus dans l'impact budgétaire.

M^{me} PARIS.- Tout à fait. C'est entendu. Y a-t-il d'autres commentaires sur cet avis ?

, merci pour la présentation. Je veux juste aller dans le sens de ce que tu as exprimé, dans la même direction. Pour en revenir à la question de l'incertitude et à l'utilisation de « mésestimée », compte tenu des limites méthodologiques que tu as identifiées, il y a vraisemblablement une mésestimation, plus qu'un effet qui est mal connu. Le problème, si l'on dit que c'est mal connu, c'est que l'on va avoir une part de l'incertitude qui n'est pas quantifiable, et nous en reviendrons à l'avis de la fin de matinée. Dans ce cas, on dit que l'on ne peut pas accorder une réserve majeure.

En revanche, je suis complètement d'accord avec toi, les choses vont dans le sens de quelque chose qui est mésestimé, avec toutes les limites que tu as pointées. L'effet relatif est foncièrement mésestimé. En disant qu'il est mal connu, l'industriel peut revenir après en disant « vous savez que nous n'avons pas d'essai en troisième ou quatrième ligne ». Je pense que le terme « mésestimée » colle bien avec ton argumentation, qui est solide et approfondie.

Un Chef de Projet du SEESP.- Oui, en sachant que nous sommes bien sur l'effet relatif en première ligne.

Tout à fait.

M^{me} PARIS.- C'est entendu, merci. Y a-t-il d'autres commentaires ? Je ne vois plus de mains levées. Je propose donc que nous votions, Carine, puisqu'il nous reste un tout petit peu de temps.

(Il est procédé au vote par appel nominatif.)

Résultat du vote :

Favorable : 15 voix.

M^{me} FERRETTI, pour la HAS.- Il y a 15 votants, 15 votes pour, donc l'avis est adopté.

M^{me} PARIS.- Merci beaucoup.