



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

❖ **Validation d'un avis économique - Avis économique – SKYCLARYS indiqué pour le traitement de l'ataxie de Friedreich (AF) chez les adultes et les adolescents âgés de 16 ans et plus (BIOGEN France SAS) – ECO-EFFI 744**

M^{me} CHEVREUL.- Vous avez vu que ce matin, on s'est concentré sur l'évaluation plutôt côté prévention dans les missions de la CEESP. On a fait un premier pas vers l'évaluation médico-économique en répondant aux remarques de l'industriel. Cet après-midi, il y a deux avis qui doivent être expertisés par la CEESP. Je sais que pas mal d'entre vous sont revenus vers moi en me disant ce serait bien qu'on ait une petite présentation de l'évaluation médico-économique parce qu'effectivement, il y a de différentes valences.

Ce que je vous propose, c'est qu'on commence par les avis, comme ça, on est sûr de rester dans notre timing. Une fois que la CEESP est finie, on discutera sur ce qu'on pourrait faire pour s'apporter les uns les autres, et avoir une mise à niveau de certains ou, peut-être une ouverture de chakras de certains vers l'évaluation médico-économique ou autre. D'accord.

Alexandre, je vous passe la parole, justement, pour le premier avis. Je n'ai toujours pas l'ordre du jour parce que mon ordinateur est déchargé, et je ne me souviens pas parfaitement.

On est dans une situation particulière, il est proposé pour l'instant que les gens qui sont encore à la Commission de la transparence écoutent, participent, disent ce qu'ils en pensent, mais ne participent pas à la prise d'avis au débat final et au vote. Par contre, vous n'aurez pas besoin de sortir, on ne vous demandera juste pas votre avis.

M. BEAUFILS, pour la HAS.- On mettra abstention pour être tout à fait francs dans le vote.

M^{me} CHEVREUL.- On met abstention, on ne demande pas aux suppléants de voter ?

M. BEAUFILS, pour la HAS.- Non, abstention.

M^{me} CHEVREUL.- Vous savez déjà que vous pouvez vous abstenir, mais n'hésitez pas à faire part de ce que vous avez pensé lors des échanges. On va pouvoir recommencer.

M. BEAUFILS, pour la HAS.- Oui, on a le quorum, on a la sténo, l'enregistrement est en cours. On va vous présenter en ce début d'après-midi le dossier SKYCLARYS par les cheffes de projet.

Une cheffe de projet, pour la HAS.- Bonjour à tous, nous allons vous présenter le dossier SKYCLARYS, l'omaveloxolone, dans le traitement de l'ataxie de Friedreich chez les adultes et adolescents âgés de 16 ans et plus. Sur ce projet d'avis, nous remercions chaleureusement Jennifer, qui est notre rapporteur, et notre référent méthodologique, qui devrait nous rejoindre.

Pour cette présentation, ce n'est peut-être pas le dossier le plus facile pour votre premier dossier, désolée, nous avons essayé de faire au maximum pour rendre la présentation fluide, mais le

dossier n'est pas évident. Désolée par avance pour cette complexité, ce n'est pas le cas de tous les dossiers, on vous rassure.

Tout d'abord, sur le contexte de cette évaluation, c'est une primo-inscription. Il y a une AMM centralisée obtenue en février 2024 et l'omaveloxolone a reçu une désignation du médicament orphelin en 2018, dans le traitement de l'ataxie de Friedreich. En termes de revendication du laboratoire, c'est un service médical rendu important qui a été revendiqué, et une ASMR III par rapport à la prise en charge actuelle des patients atteints de l'ataxie de Friedreich. On a une indemnité dans le cadre de l'accès précoce qui est de 23 200 euros par boîte de 90 gélules.

Quelques mots sur la pathologie. L'ataxie de Friedreich est une maladie génétique rare neurodégénérative qui va être d'évolution lente liée à une atteinte du cervelet et des voies sensitives profondes. La maladie va se manifester principalement par des troubles de la marche, de l'équilibre, de la coordination des membres supérieurs, des difficultés à réaliser ensuite des activités dans la vie quotidienne des patients.

À ce jour, il n'y a pas de traitement spécifique de la pathologie. C'est plutôt une prise en charge globale et précoce des patients pour préserver au maximum les capacités fonctionnelles, l'autonomie et éviter toutes les complications qui sont de différentes sortes neurologiques, ostéoarticulaires, cardiaques, etc.

La prise en charge des patients va reposer sur un suivi multidisciplinaire. Il va être coordonné par le médecin traitant, le neuro ou neuropédiatre et le neurogénéticien référent.

Dans le cadre de ce dossier, il y a une contribution d'une association de patients, l'Association française de l'ataxie de Friedreich. Dans le projet d'avis, vous retrouvez un résumé de la contribution patient qui est proposé au début du document. On pourra revenir dessus tout à l'heure si vous le souhaitez. Il a été souligné l'impact sur la vie sociale, affective, scolaire et professionnelle des patients de cette pathologie et le caractère évolutif et incurable avec une très grande dépendance des patients et une réduction de l'espérance de vie.

Pour continuer sur les revendications de l'industriel dans ce dossier, il y a un RDCR de près de 4,7 millions par QALY qui est revendiqué dans la stratégie de prise en charge par l'omaveloxolone plus meilleurs soins de support par rapport aux meilleurs soins de support, avec un prix environ de [REDACTED] euros pour de boîte de 90 gélules dans la modélisation. L'impact budgétaire est de 545 millions d'euros sur trois ans. Le chiffre d'affaires prévisionnel en deuxième année de commercialisation revendiqué par le laboratoire est d'environ [REDACTED] d'euros. Le laboratoire ne revendique pas d'incidence sur l'organisation des soins, les pratiques professionnelles et les conditions de prise en charge des malades.

Quelques mots sur les données cliniques. L'étude qui est proposée dans le cadre de ce dossier, c'est l'essai clinique MOXle, la partie 2. C'est une étude de phase II, de supériorité, multicentrique, randomisée en double aveugle en groupes parallèles, comparative par rapport au placebo. Partie 2 puisqu'il y a deux parties dans cette étude, une première partie qui était pour déterminer la dose minimale efficace et une deuxième partie, MOXle 2, qui est utilisée ensuite dans le dossier. C'est une étude réalisée chez 103 patients. L'objectif était de démontrer la supériorité de l'omaveloxolone par rapport au placebo sur une durée de 48 semaines.

Peut-être un point important sur les critères d'inclusion, ce sont des patients qui étaient âgés entre 16 et 40 ans qui étaient inclus, avec un score mFARS compris entre 20 et 80. Le score mFARS permet d'évaluer l'aspect neurologique des patients, c'est une échelle qui peut monter jusqu'à 93 points. Plus le score est élevé, plus la maladie évolue.

Le critère de jugement principal était la variation du score mFARS entre l'inclusion et 48 semaines dans le groupe omaveloxolone par rapport au placebo. Le traitement par omaveloxolone a significativement amélioré le score mFARS par rapport au placebo avec une différence de 2,4 points sur cette échelle.

Un intervenant.- J'ai une question par rapport à ce résultat, a-t-on une idée de la différence minimale qui est significative sur cette échelle ? Cela a été évalué à un moment donné ? Parce que vous avez dit que l'étendue c'était 90 points ?

Une cheffe de projet, pour la HAS.- 93 points.

M. AMABILE.- Là on voit, en plus c'est une figure de l'article ? Parce que la figure est trompeuse. On a l'impression que c'est une différence qui est très importante entre les deux groupes, alors que 2,4 points sur 93 points, ça remet en perspective un peu.

M. BEAUFILS, pour la HAS.- C'est une excellente question. Les chefs de projet ont des échanges avec les experts, notamment ceux qui étaient passés en CT. Des discussions ont d'ailleurs eu lieu en CT exactement sur ce sujet. C'est statistiquement significatif, effectivement, après la question de savoir est-ce cliniquement pertinent, c'est là que c'est beaucoup plus compliqué à définir. Il n'y a pas de valeur dans la littérature qui dise à partir de cette valeur, c'est cliniquement pertinent. D'après le laboratoire ou des experts du laboratoire, c'est cliniquement pertinent et quand on écoute les experts, on se pose vraiment la question de cette différence, comme vous le dites, sur une échelle de 90 points. Ce n'est pas évident. C'est statistiquement significatif, mais cliniquement pertinent, ce n'est pas évident de le dire.

Une intervenante.- J'avais une autre petite question. C'est un essai clinique qui a été fait uniquement en France ?

Une cheffe de projet, pour la HAS.- Il n'y a pas de patients français dans l'étude clinique. Il y a 11 centres, 5 pays : US, Australie, Autriche, Italie et Royaume-Uni, mais pas de patients français.

Pour le suivi de cette étude, il y a un suivi en ouvert des patients qui a été réalisé. L'objectif, c'était de regarder la tolérance à long terme de l'omaveloxolone. C'est l'étude MOXle Extension qui est proposée.

Dans le package de données, on a une comparaison qui a été faite avec appariement sur score de propension des données individuelles qui a été réalisé avec l'étude MOXle Extension et une cohorte historique qui s'appelle FA-COMS dont le but était d'estimer l'efficacité de l'omaveloxolone sur une durée de trois ans. Cela va être repris par la cheffe de projet sur cette comparaison avec Appariement.

Il y a une autre question.

M. SCHWARZINGER.- Avec beaucoup de retard, mais si on peut poser des questions au fur et à mesure, c'est super. Comme c'est ma première session, j'ai juste une question sur l'industriel qui ne revendique pas d'impact sur l'organisation des soins, etc. À partir de combien de milliards on considère qu'il y a un impact ?

M. BEAUFILS, pour la HAS.- Oui, c'est vrai que cela peut paraître un peu discordant. Dans la loi, ce qui fait que l'on est éligible à une évaluation économique, qu'on doit soumettre une évaluation économique, c'est effectivement si on a un chiffre d'affaires revendiqué supérieur à 20 millions d'euros ou si on revendique une incidence sur l'organisation des soins, des pratiques professionnelles, les conditions de prise en charge des malades. Mais ce n'est pas clairement explicité ce qui rentre dans la définition « organisation des soins, pratiques professionnelles ou conditions de prise en charge des malades ». C'est-à-dire qu'on peut se rendre compte qu'il y a des choses qui peuvent être un peu... ce n'est pas très logique.

Dans les faits, les produits sont généralement évalués parce qu'ils revendiquent un chiffre d'affaires supérieur à 20 millions d'euros. Ils ne revendiquent pas tant des incidences sur l'organisation des soins, les pratiques professionnelles et les conditions de prise en charge des malades. Il n'y a pas de lien avec le chiffre d'affaires revendiqué, bien qu'il pourrait y avoir un lien.

M^{me} CHEVREUL.- Ce n'est pas une des raisons pour lesquelles ils passent en CEESP.

M. BEAUFILS, pour la HAS.- Non, pas du tout. C'est une case à cocher. Je vous le dis comme ça.

M. SCHWARZINGER.- C'est une case à cocher.

M. BEAUFILS, pour la HAS.- C'est une case à cocher, exactement. Je vous le dis comme ça.

M^{me} CHEVREUL.- Peut-être que ce qu'il faudrait dire c'est quand c'est passé en CT.

M. BEAUFILS, pour la HAS.- En juillet.

Une cheffe de projet, pour la HAS.- Il y a eu l'audition la semaine dernière.

M. BEAUFILS, pour la HAS.- C'est passé plusieurs fois.

M^{me} CHEVREUL.- Il n'y a pas de SMR et d'ASMR ?

Une cheffe de projet, pour la HAS.- Si.

M. BEAUFILS, pour la HAS.- C'est modéré, V.

M^{me} CHEVREUL.- On s'est dit que ce qui serait peut-être intéressant c'était de savoir ce qui avait été discuté et s'il y avait eu des points d'achoppement sur l'évaluation de l'efficacité ou pas, et où avait porté la discussion.

M. BEAUFILS, pour la HAS.- Le but aujourd'hui, nous n'avons pas préparé de vous faire un compte-rendu des discussions de la CT, mais le vote d'un SMR modéré, d'une ASMR V qui traduit une absence d'amélioration par rapport à la stratégie de prise en charge, je ne vais pas dans les détails, mais ça va dans le sens que la CT était assez critique par rapport aux résultats démontrés ou pas démontrés de ce produit.

M^{me} CHEVREUL.- Merci.

Une cheffe de projet, pour la HAS.- On vous propose de passer à la partie de l'analyse de l'efficacité. Tout d'abord, l'objectif et les choix structurants. L'objectif de cette analyse est d'évaluer l'efficacité de l'omaveloxolone dans le traitement de l'ataxie de Friedreich chez les adultes et adolescents âgés de 16 ans et plus, en considérant toutes les options thérapeutiques.

Dans ce dossier, nous avons une analyse coût-utilité qui a été fournie uniquement et pas d'analyse coût-efficacité. Pourquoi ? Parce que le laboratoire ne revendique pas d'impact du traitement sur la mortalité des patients. C'est pour cette raison qu'on n'a qu'une analyse coût-utilité qui est proposée.

La perspective de l'analyse est une perspective restreinte au système de santé. L'horizon temporel, c'est une durée limitée à 20 ans. Le taux d'actualisation est de 2,5 % jusqu'à 30 ans, puis 1,5 %, mais comme on est limité à 20 ans, ça n'entre pas en ligne de compte. La population d'analyse, ce sont les patients adultes et adolescents âgés de 16 ans et plus atteints de l'ataxie de Friedreich. Les comparateurs de cette analyse, ce sont les meilleurs soins de support qui vont être définis comme les traitements et soins associés à la prise en charge des complications de l'ataxie de Friedreich, c'est-à-dire complications neurologiques, cardiaques, ostéoarticulaires, etc. Le comparateur qui n'est pas retenu, c'est l'idébénone qui nous paraissait acceptable à ne pas retenir.

En termes d'analyse qui est proposée dans le dossier, le fait de ne pas proposer une analyse coût-efficacité nous paraissait acceptable dans ce contexte. L'horizon temporel de 20 ans, on a rapporté qu'il pouvait être discutable au regard de certaines incertitudes qu'on pourrait avoir, notamment sur la durée de traitement et le maintien de l'effet et l'absence d'impact de mortalité sur les patients. Néanmoins, cela nous a paru plutôt acceptable au regard de la pathologie d'évolution lente.

Je passe la parole sur la population simulée.

Une cheffe de projet, pour la HAS.- La population simulée, dans le modèle, correspond à des patients adultes et adolescents âgés de 16 ans et plus et atteints d'ataxie de Friedreich. Elle correspond finalement à la population demandée au remboursement. Cette population simulée est estimée en partie sur les données disponibles dans plusieurs études. Nous avons l'étude Pousset et al., qui est une étude française rétrospective conduite sur une centaine de patients. Nous avons également l'étude Reetz et al., qui est un registre européen nommé EFACTS. Enfin, nous avons l'étude Rummy et al., qui est une cohorte internationale qui comprend environ 1 000 patients et qui a à peu près 18 ans de suivi.

Pourquoi ai-je dit « en partie » ? C'est parce qu'en réalité, l'industriel propose dans ce dossier quelque chose d'assez atypique, à savoir une simulation d'une population simulée, conduite en plusieurs étapes. La première étape proposée par l'industriel est de simuler une cohorte représentative de l'histoire naturelle de la maladie en faisant un tirage aléatoire des caractéristiques des patients sur la base d'une équation qui permet d'estimer l'évolution du score mFARS en fonction de l'âge.

Cette équation est très importante, je vais vous la présenter dans les diapositives suivantes, mais elle permet d'estimer le paramètre clé de notre modèle, qui est l'évolution du score mFARS. Elle fait une relation entre cette évolution du score et l'âge des patients en considérant plusieurs caractéristiques. Cette équation va être utilisée pour estimer notre cohorte représentative de l'histoire naturelle et l'évolution des différentes caractéristiques cliniques de nos patients.

Dans une seconde étape, l'industriel propose de sélectionner un échantillon de cette première cohorte en retenant uniquement les patients qui ont un âge à l'initiation du traitement de 16 ans afin de correspondre à la population demandée au remboursement et également en ajustant les caractéristiques cliniques qui ont été estimées sur la durée moyenne de la maladie qui est issue de l'étude française afin d'être plus représentative de la population qu'on aura en pratique clinique en France.

Une première analyse sur ces premiers points, c'est qu'effectivement concernant la première étape, comme les caractéristiques simulées sont issues de cette équation d'évolution du score

mFARS, elle reflète toutes les limites de cette dernière. Ici, nous proposons de ne pas mettre une réserve à cet endroit dans le projet d'avis, puisque nous allons rediscuter ce point plus loin au moment où nous allons présenter cette équation du paramètre mFARS. Désolée, c'est très lié, donc on va faire pas mal de renvois sur les premières parties.

Sur la deuxième étape d'estimation de la population simulée, il nous semble que le choix d'ajuster la cohorte sur la durée moyenne de la maladie observée en France était un choix intéressant, puisque cela permettait d'être plus représentatif et de correspondre davantage à l'indication demandée au remboursement. Toutefois, cela peut avoir un impact, puisque finalement, on va restreindre le périmètre des caractéristiques de nos patients et on pourrait avoir un impact potentiel sur les paramètres qui sont dépendants de l'âge. En effet, on va le voir notamment grâce à cette équation. L'âge a un impact dans l'évolution de la maladie en fonction de l'apparition des symptômes, par exemple. Finalement, cet ajustement sur la durée moyenne pourrait avoir un impact sur notre population simulée.

Toutefois, sur ce point, nous avons tout de même une analyse de sensibilité qui est conduite, puisqu'il est proposé d'inclure en analyse de sensibilité les caractéristiques des patients de l'essai MOXle-2 et cela entraîne une diminution de 16 % du RDCR.

Finalement, ici, vous voyez les caractéristiques moyennes de la population simulée qui sont issues pour la durée moyenne de la maladie de l'étude Pousset et al., sinon pour les autres caractéristiques issues en partie de l'équation de modélisation.

Ensuite, en termes d'exercice d'analyse de la représentativité, l'industriel proposait dans son dossier de comparer les caractéristiques des patients de l'essai clinique MOXle partie 2 à celles des études observationnelles Pousset et al., et Reetz et al. Sur ce point, nous avons plusieurs analyses critiques qui sont proposées.

En effet dans un premier temps concernant les caractéristiques de la population simulée puisqu'elles sont tirées au sort, notre premier point d'analyse est que, finalement, ces caractéristiques ne correspondent pas aux patients inclus dans l'essai clinique, mais également ne correspondent pas non plus aux patients inclus dans les études observationnelles mobilisées.

Lorsque l'on fait l'analyse de la représentativité proposée par l'industriel, finalement, on ne compare pas réellement notre population simulée dans notre modèle aux données qui vont être mobilisées dans l'analyse économique, que ce soit l'essai clinique ou les études observationnelles. Finalement, en l'absence de cette analyse conduite de manière cohérente avec les bonnes données qui ont été utilisées dans le modèle, aujourd'hui, il nous semble que nous ne sommes pas en mesure de dire que la transposabilité de la population simulée peut être assurée afin d'être

représentatif de la population susceptible d'être traitée en France ou des populations qui sont issues des études mobilisées.

Au regard de l'ensemble de ces éléments, nous proposons une première réserve importante qui est sur « La transposabilité de la population simulée aux populations issues des études mobilisées et à la population susceptible d'être traitée en France, non assurée, en raison de l'analyse de la représentativité de la population simulée qui est non cohérente. »

Concernant la structure du modèle, ici, l'industriel a fait le choix de proposer un modèle de microsimulation à temps continu fondé sur l'évolution continue du score mFARS dans le temps. Il justifie ce choix puisque l'atteinte neurologique est la principale manifestation clinique et la première source de handicap pour les patients. Il justifie également ce choix puisque, selon lui, ce modèle permet de simuler de manière continue l'évolution du score mFARS en intégrant ainsi la variabilité individuelle des caractéristiques des patients.

Selon nous, le choix d'un modèle de microsimulation est acceptable. En effet, de la compréhension de la pathologie, les patients peuvent être pris en charge à différents stades dans leur maladie. Il existe aussi une forte hétérogénéité individuelle de l'expression clinique en fonction du score mFARS. Pour un même score mFARS, on peut avoir une expression clinique différente entre les patients. Ce modèle semble adéquat.

Toutefois, il pose quelques limites et là, nous verrons cela également dans les sections ultérieures. Mais il nécessite de considérer que la qualité de vie évolue également de manière continue et il nécessite pour l'intégration des coûts de définir des états de santé afin d'appliquer pour chaque état de santé un coût de prise en charge. Ces points seront décrits par la cheffe de projet plus tard.

Je propose maintenant de passer à un point clé du modèle qui est les paramètres d'évolution du score mFARS. Pour estimer ces paramètres d'évolution du score mFARS, l'industriel propose de se baser sur des hypothèses cliniques qui vont être présentées ensuite d'utiliser la calibration d'une équation à partir des données observées dans la cohorte historique FA-COMS. Trois hypothèses sont proposées.

La première hypothèse est qu'il existe une corrélation entre la sévérité de la maladie, qui peut être définie notamment par la longueur du triplé GAA1, et la vitesse de progression du score mFARS.

La deuxième hypothèse proposée par l'industriel est que la vitesse de progression du score mFARS est corrélée à l'âge d'apparition des symptômes.

La troisième hypothèse est que la progression du score mFARS est linéaire avec un point d'inflexion lors de la perte d'ambulation.

Ces trois hypothèses reposent sur des données disponibles dans la littérature, notamment les données de la cohorte FA-COMS, celle de la cohorte EFACTS, enfin de l'étude Patel dont on va discuter par la suite.

Mais avant, je continue sur la méthode d'estimation d'évolution du score mFARS. Effectivement, à partir de ces hypothèses, il est proposé de traduire ces hypothèses cliniques en fonctions mathématiques avec l'équation que vous avez présentée sur la diapositive qui permet finalement d'estimer l'âge pour un score mFARS en prenant en considération l'âge à l'initiation de la maladie, le score mFARS à la perte d'ambulation et les vitesses de progression de score mFARS selon le statut ambulatorio. C'est pour cela que vous avez ici deux équations. Cela permet de refléter ce qui est proposé dans le modèle. Cette évolution continue avec un impact important de l'âge sur l'évolution du score mFARS. Par la suite, cette équation a également été calibrée sur les données de la cohorte FA-COMS pour estimer justement la moyenne et l'écart-type de chacun de ces paramètres.

Maintenant, je propose de passer à l'analyse critique que nous vous proposons sur ces différents points. Comme vous avez pu le voir, notre premier point est qu'effectivement, les hypothèses cliniques de l'industriel vont structurer l'ensemble des paramètres. Par conséquent, nous sommes allés vérifier la plausibilité de chacune de ces hypothèses.

L'hypothèse 1, selon laquelle il existe une corrélation de la sévérité de la pathologie à la vitesse de progression du score mFARS, est plus ou moins discutable. En effet, dans l'étude Patel, selon le recours à différents modèles statistiques, il existe une non-convergence entre les différents modèles, qui met en évidence que la corrélation est discutable entre le score mFARS et la sévérité de la maladie.

Concernant la deuxième hypothèse, la vitesse de progression du score mFARS serait corrélée à l'âge d'apparition des symptômes. Cette hypothèse est effectivement soutenue par les données disponibles dans l'étude Patel. Toutefois, l'industriel avançait également que les données de l'étude Reetz pouvaient soutenir cette hypothèse, sauf que les paragraphes et graphiques mentionnés par l'industriel étaient non pertinents pour soutenir cette hypothèse.

Enfin la troisième hypothèse, selon laquelle la progression du score mFARS est linéaire avec un point d'inflexion lors de la perte d'ambulation, repose uniquement sur des données observationnelles.

Ici, aucune des études mentionnées ne démontre qu'il existe une dynamique différente selon le statut ambulatoire de manière statistiquement significative. Par conséquent, il ressort que cette hypothèse semble avoir une forte incertitude et n'est pas explorée en analyse de sensibilité dans le dossier.

Finalement, sur cette première étape d'estimation des paramètres d'évolution du score mFARS, il persiste une forte incertitude puisque ces paramètres sont estimés sur la base d'une équation qui n'est ni établie ni validée dans les littératures, qui correspond uniquement à la traduction d'hypothèses cliniques en fonctions mathématiques. L'impact de ce choix est supposé important sur les résultats. En effet, l'estimation de ces paramètres est fortement liée dans le modèle à l'estimation des années de vie.

Au regard de l'ensemble de ces éléments, nous proposons sur cette première étape une réserve importante qui est la suivante : « Méthode d'estimation des paramètres d'évolution du score mFARS qui nécessite le recours à une équation non établie et non validée dans la littérature, générant une forte incertitude quant à l'évolution des patients dans le modèle et qui est non explorée en analyse de sensibilité. »

Par la suite, l'intégration de ces fameux paramètres d'évolution du score mFARS dans le modèle. Finalement, on l'a vu un peu lorsque je vous ai présenté la population simulée. À travers cette équation, nous allons avoir une simulation des caractéristiques des patients qui vont représenter l'évolution naturelle de la maladie avec un tirage aléatoire de trois paramètres qui sont la vitesse de progression du score mFARS, selon le statut ambulatoire ou non, le score mFARS au diagnostic et le score mFARS à la perte d'ambulation. Ces paramètres vont être tirés de manière aléatoire selon une distribution gamma de moyenne et d'écart-type qui aura été estimée par la méthode de calibration qui était conduite sur les données de la cohorte FA-COMS. N'hésitez pas si vous avez des questions.

M^{me} CHEVREUL.- Ce qu'on va faire, si tu es d'accord, c'est que l'on finisse celui-là, et on voit après. On fait deux réserves à la fois.

Une cheffe de projet, pour la HAS.- Parfait. Du coup, sur notre analyse critique, comme on l'a dit tout à l'heure, les caractéristiques simulées des patients de la cohorte sont finalement issues de cette équation et ne proviennent pas directement des données cliniques et des données observationnelles. Ce point a déjà fait l'objet d'une analyse et d'une réserve dans la partie population simulée.

Toutefois, nous avons également relevé un élément supplémentaire. Finalement, le choix de la distribution pour l'estimation des paramètres, ici la distribution gamma, n'a pas été justifié. L'industriel précise uniquement que ces distributions permettent de refléter l'histoire naturelle

de la maladie, mais sans aucun détail supplémentaire. Au regard de ces éléments, nous proposons une réserve mineure qui est la suivante : « Manque de justification quant aux choix des fonctions de distributions pour l'estimation des paramètres d'évolution du score mFARS. »

M^{me} CHEVREUL.- Ce que je vais vous proposer, si vous êtes d'accord, c'est qu'on ait peut-être la possibilité d'avoir un échange réserve par réserve, justement pour savoir si vous avez des questions, si vous voulez des spécifications par rapport, même pour certains qui sont moins aguerris à l'évaluation médico-économique, la signification, ce qu'on est en train de chercher. C'est l'occasion aussi d'en faire un exercice pédagogique, pour certains et d'autres qu'ils puissent avoir des remarques qui soient plus spécialisées. Vous n'êtes pas obligés d'avoir des questions. C'est juste pour donner de l'occasion à chacun de pouvoir le faire. Certains ont-ils des questions sur les paramètres d'évolution du score mFARS ?

Sandrine.

M^{me} LOUBIERE.- Peut-être pas sur ces paramètres, mais sur la représentativité de la population. C'est vrai que c'était complexe, mais je me posais la question, vu que vous avez évoqué, il me semble, qu'il y avait des registres, ou je me trompe, la possibilité de tester la représentativité de leur population simulée avec ces registres n'était pas possible ?

Une cheffe de projet, pour la HAS.- Ce qu'ils nous proposent aujourd'hui, et ce qui nous pose souci dans l'analyse de la représentativité, c'est qu'ils nous proposent de comparer des populations qui, finalement, ne sont pas celles incluses dans le modèle. Effectivement, ils auraient été en mesure de nous proposer une telle analyse. Après, ce qui est un peu particulier, c'est que, pour avoir mis les mains dans le modèle, comme toutes les caractéristiques sont dépendantes, liées à travers l'équation et tirées au sort pour chaque patient, on a des cohortes fictives dans le modèle, dont ils nous ont proposé la moyenne pour certains des paramètres. Là, vous voyez, on en a juste trois, finalement. Ils ne nous ont pas fait l'exercice de nous présenter la totalité des caractéristiques des patients sur tous les paramètres. On a juste ce petit tableau qui est tout de même très léger pour conduire une analyse de représentativité.

M^{me} CHEVREUL.- Merci. Une question d'Antoine, avant Gaëtan.

M. BENARD.- Je ne suis pas sûr d'avoir bien compris la méthode de simulation qui a été effectuée. Je ne suis pas sûr que ce soit compréhensible. Vous avez dit modèle à temps continu. Je connais les modèles à événements discrets, c'est-à-dire qu'on essaie de prédire la survenue d'événements. Ce n'est pas le même type de modèle ?

Une cheffe de projet, pour la HAS.- Non.

M. BENARD.- Pouvez-vous nous expliquer ce que c'est un modèle à temps continu ?

Une cheffe de projet, pour la HAS.- Oui. Voulez-vous que je revienne d'abord sur la population simulée, peut-être ?

M. BENARD.- Non.

Une cheffe de projet, pour la HAS.- Je peux commencer et si l'autre chef de projet veut compléter la réponse, n'hésite pas. Là, on n'est pas sur un modèle de microsimulation d'événements discrets. On est bien sur un modèle de microsimulation à temps continu, donc on a une évolution continue de nos paramètres, de la totalité de nos paramètres, de notre efficacité, mais également de notre qualité de vie. On modélise cette évolution continue et sans arrêt du score mFARS. C'est pour cela qu'on a cette équation, qui est une équation de droite, qui est présentée ici sur le graphique, comme vous le voyez. Vous avez le temps et l'évolution du score mFARS dans le temps. Comme on l'a dit, elle est représentée de manière linéaire grâce à notre équation, avec un point d'inflexion quand on a la perte d'ambulation, on est bien sur quelque chose de continu.

Cela a des avantages, mais cela va poser plusieurs limites pour l'intégration de la qualité de vie ou même pour l'intégration des coûts, puisqu'on peut difficilement considérer que les coûts sont des variables continues.

Je ne sais pas, le chef de projet veut rajouter un mot sur le modèle.

Un chef de projet, pour la HAS.- Je viens juste d'arriver. Bonjour et bienvenue d'abord à tout le monde. Je me suis déjà présenté la dernière fois. Historiquement, les modèles de microsimulation ont commencé avec les modèles avec les ingénieurs, les aéroports et tout, avec le temps continu. Après, les modèles à choix discret, c'est bizarre dans le domaine de la santé aujourd'hui de ne pas proposer un modèle à choix discret à des instants précis, prédéfinis.

Ils ont choisi une estimation d'un modèle individuel de microsimulation individuelle où, à chaque instant, bien sûr, c'est discret au sens relatif. Il y a des instants dans le temps. De cette manière, ils ont pu calculer le QALY comme étant une aire de la surface, comme tu dis l'intégral de la formule usuelle. Cette spécification aux formes fonctionnelles est liée à leur façon d'interpréter l'histoire de la pathologie de manière dynamique. C'est vrai, je suis complètement d'accord avec vous et on a posé la question. D'ailleurs, les collègues vont venir en parlant des coûts par la suite. Ils proposent par la suite, faute de mieux, parce que c'est la réalité, on ne peut pas définir les coûts de manière continue, cela n'a pas de sens. Ils les ont donc définis de manière catégorielle, comme s'ils correspondent à des événements, à des modèles à choix discret à des moments donnés. Effectivement, ils auraient pu peut-être choisir un modèle à choix discret, mais ils ont suivi leur logique.

M. BENARD.- Ce que je comprends, c'est qu'ils ont essayé d'estimer à chaque temps quelle est la valeur du score mFARS, avec une hypothèse de continuité entre deux points dans la phase ambulatoire, et un autre point dans la phase non ambulatoire.

Une cheffe de projet, pour la HAS.- Tout à fait.

M. BENARD.- Ils utilisent un certain nombre de variables pour estimer ce score mFARS à chaque temps, c'est ça ?

Une cheffe de projet, pour la HAS.- Oui. Une fois qu'ils ont identifié où le patient arrive sur ce graphique, une fois qu'il est mis sur la courbe, après il suit l'évolution qui est proposée.

M. BENARD.- C'est essentiellement dépendant de la mFARS.

Une cheffe de projet, pour la HAS.- En majeure partie.

M^{me} CHEVREUL.- Gaëtan avait également une question.

Une cheffe de projet, pour la HAS.- Oui, il y avait la question sur la population simulée. C'était sur le modèle ?

M. DUPORT.- Si vous voulez répondre sur la population, parce que j'ai un point à apporter sur la population aussi. Je voulais rebondir sur l'histoire de la population et sur l'histoire de l'organisation des soins. Il faut se mettre dans la situation de l'ataxie de Friedreich. C'est une maladie extrêmement grave, dégénérative, sur laquelle on fait un diagnostic. On sait faire le diagnostic, mais sur lequel on n'a aucun traitement à proposer. Que font les patients dans cette situation et les familles ? Pour la moitié d'entre elles, en discutant avec l'Association française d'ataxie de Friedreich, ils nous expliquent que la moitié des familles fuit les centres de ressources et compétences puisqu'il n'y a aucun traitement. Quand on parle de population, pourquoi on a pris la population des registres ? Parce qu'ils sont sûrement inclus au moment du diagnostic après, pour la moitié, on ne sait pas où ils sont.

Avec l'accès précoce qui a été mis en place il y a plusieurs mois, six mois si je ne m'abuse, certaines familles sont revenues puisque maintenant, il existe quelque chose. On est en train de voir en détail ce qu'il en est ici, mais du coup un surcoût est engendré. Quand on parle d'organisation de soins, on a des patients qu'on ne voyait plus du tout. Puis, ils arrivent et du jour au lendemain, ils demandent à avoir accès à cette molécule et ils demandent aussi à être suivis, puisqu'à partir du moment où vous avez un traitement, vous avez envie d'être suivi par des professionnels de santé.

Les standards de soins qui sont mis en place pour ces patients, ce sont surtout des soins qui sont des soins type paramédicaux, le kinésithérapeute, l'orthophonie, des soins sur lesquels on peut renouveler en permanence avec un médecin généraliste sans être justement dans les centres de

ressources et compétences. C'était juste mon point sur la population pour expliquer pourquoi les registres ne permettent pas d'avoir des informations claires.

M^{me} CHEVREUL.- Merci.

Une cheffe de projet, pour la HAS.- Merci pour ces éléments. Effectivement, ce qui est particulier ici dans ce dossier, sur l'estimation de la population simulée, c'est que finalement, l'industriel fait le choix de ne prendre en considération ni les données de l'essai clinique, ni les données d'études observationnelles. Usuellement, on prend les données de l'essai clinique ou des données d'études observationnelles, et on utilise les caractéristiques de ces populations pour refléter les caractéristiques de notre population simulée. Ici, le choix de l'industriel a été de simuler de manière fictive une cohorte de patients.

Le problème de la transposabilité et de la question de la transposabilité se posent, puisque cette cohorte ne reflète ni les données mobilisées dans l'étude, ni complètement les données qu'on aurait pu observer en France et donc les patients susceptibles d'être traités.

Les étapes de la constitution de la population simulée sont un peu complexes et longues, mais effectivement, l'industriel va simuler cette cohorte fictive en utilisant les paramètres de l'équation qui permettent de refléter, selon lui, l'évolution du score mFARS, ce que je vous ai présenté. Sur ces paramètres, il va tirer de manière aléatoire, en suivant une distribution gamma des paramètres pour chacun des patients simulés pour créer cette cohorte fictive.

Ensuite, il va proposer dans une deuxième étape de faire des ajustements pour essayer de faire en sorte que cette cohorte fictive reflète au mieux les patients attendus dans l'indication et en France, en faisant un ajustement sur l'âge d'initiation du traitement, puisque le traitement en France va être donné aux patients qui ont 16 ans et plus, comme l'indication le précise, et en ajustant également sur la durée moyenne, puisque finalement, comme on l'a vu, l'âge au début du traitement, l'âge au diagnostic sont des paramètres qui vont avoir un impact dans la modélisation de l'évolution du score mFARS des patients.

Ici, en ajustant sur la durée moyenne de la maladie en France, ça permet de refléter un peu mieux, d'avoir une cohorte qui est plus représentative, puisque l'on aura des patients qui ont en moyenne une durée de la maladie depuis 15 ans. Je ne sais pas si c'est plus clair et si vous avez d'autres questions.

Une intervenante.- Merci beaucoup, parce que vous nous avez très bien expliqué la méthode de simulation. Ce n'est pas évident, mais on a tous les éléments, et merci pour ces compléments. Pouvez-vous juste revenir sur la raison pour laquelle ils n'ont pas utilisé les données d'essai clinique ? Pardonnez-moi, vous avez été peut-être rapide, ou je n'ai pas entendu.

Une cheffe de projet, pour la HAS.- On leur a demandé en échange technique. On a une analyse de sensibilité avec ces caractéristiques, mais le choix de l'avoir présentée en analyse de référence n'a pas été justifié par l'industriel. On n'a pas d'éléments sur ce point.

M. SERRIER.- J'essaie de m'accrocher, mais c'est un peu les mêmes questions. Il y a un truc que je ne suis pas sûr de comprendre. Soit on a une population, celle de l'essai ou autre, dont on sait qu'elle est représentative de la France, et dans ce cas, on sait que leur population n'est pas représentative de la population. Soit on n'a pas de population vraiment représentative de la France, et dans ce cas, on a un doute sur la représentativité de la population qu'ils ont simulée. J'ai un peu de mal à voir dans quelle des situations on se trouve finalement. A-t-on vraiment un référentiel ? Est-on capable de dire que leur population n'est pas représentative des patients français ? Ou le comparateur à qui on se réfère, est-il incertain ?

Une cheffe de projet, pour la HAS.- Honnêtement, il y a une étude rétrospective qui a été conduite en France avec des caractéristiques de patients. L'industriel n'a pas choisi d'avoir recours à cette étude. Comme je le disais, l'analyse de représentativité qui a été faite aujourd'hui est non cohérente, puisqu'on compare les données de l'essai clinique aux études observationnelles européennes et françaises. Maintenant, nous avons une cohorte qui est construite de manière fictive. La justification de l'industriel pour avoir fait ce choix n'a pas été apportée en réponse aux questions pendant l'échange technique.

Pendant l'échange technique, on a aussi posé beaucoup de questions sur comment avait été finalement construite cette population simulée fictive, parce que c'était tout de même un peu long à démêler. On n'a pas eu d'éléments de réponse sur pourquoi ils ont agi comme ça. Nous considérons que la transposabilité est non assurée, puisque finalement, nous n'avons pas l'analyse de représentativité adéquate. Et comme je le disais précédemment, nous n'avons même pas toutes les caractéristiques moyennes de la population simulée qui sont présentées dans le rapport technique.

Une cheffe de projet, pour la HAS.- Il y a juste un petit détail sur la population simulée de l'étude clinique. Les patients sont globalement tout de même plus jeunes et à un stade moins avancé de la maladie. Il y a probablement aussi une explication là-dessus en disant que dans l'essai, les patients sont plus jeunes et à un stade moins avancé.

M. SCHWARZINGER.- Peut-être pourriez-vous dans les commissions suivantes envoyer les papiers principaux. J'ai téléchargé le papier du fameux Pousset de 2015. Un élément de réponse, c'est qu'il n'y a pas de score mFARS qui a été mesuré dans l'étude.

Une cheffe de projet, pour la HAS.- Sur ce point, c'est vrai qu'on ne vous a pas présenté, mais en Europe, on n'utilise pas le score mFARS. On utilise l'échelle SARA. C'est quelque chose qui va arriver dans la suite de la présentation.

Maintenant, il existe une corrélation entre le score mFARS et le score SARA qui est publié et qui est suffisamment robuste pour être considérée comme acceptable. C'est vrai que ce problème de traduction SARA-mFARS n'en est pas un dans le dossier, puisqu'on fait souvent, vous allez voir, le parallèle en passant de l'un des scores à l'autre sans trop de problématiques, puisqu'on peut avoir confiance en cette corrélation.

Une intervenante.- Pardonne-moi, à quel endroit, à quelle page de l'avis présentez-vous l'essai clinique pour qu'on puisse rentrer un peu dedans ?

Une cheffe de projet, pour la HAS.- L'essai clinique, en général, est présenté de manière très succincte. C'est dans le rapport technique de l'industriel.

M^{me} CHEVREUL.- Pour le coup, c'est pour cela que ce n'est pas un bon exercice sur les nouvelles relations entre nous. C'est comme tout à l'heure, ce que je demandais. On est en train d'en parler entre les différentes commissions, sur comment on fait un retour justement sur ce qui s'est discuté en CT avant d'arriver à la CEESP, sur la validité des preuves qui ont été données sur l'efficacité ou autre, pour qu'on reparte exactement de la même chose et qu'il y ait une succession. Cela ne se faisait pas jusque-là. C'est bien que vous posiez la question avec cette nouvelle commission, parce que j'ai posé la question, et c'est pour cela qu'on est en train d'essayer d'organiser le passage.

Peut-être que cela veut dire que pour les prochaines commissions, on peut essayer de faire cette synthèse, on peut essayer d'envoyer des éléments là-dessus et voir avec les rapporteurs ou les personnes d'entre vous qui sont le plus intéressées, comment on peut organiser ce passage.

Clairement, cela n'a pas été demandé avant, mais on voit que c'est une évolution qui est tout de même assez nécessaire, d'autant que la différence, c'est qu'ici, on a des personnes qui sont très spécialisées dans l'évaluation médico-économique, mais qu'on n'a pas les méthodes qu'on a à la CT, c'est bête de perdre cette information. Surtout que cela peut justement améliorer la compréhension de certains d'entre vous.

Ce que je propose, c'est qu'effectivement, une fois qu'on a fini, qu'on puisse affiner ce que serait plus votre demande et autres en cours de route pour que ce soit le plus fluide possible, avec les chefs de projet du SEM, pour que ce soit le plus dans les deux sens, d'accord ?

M. MERCIER.- Une suggestion, c'est que dans les rapports techniques, il y ait finalement des références aux papiers qui ont été publiés. Parce qu'au niveau de la CT, cela a mis beaucoup de

temps avant d'arriver. C'est vraiment important que l'on puisse faire des références à des papiers qui ont été publiés. Par exemple, le papier qui rapporte cet essai a été publié dans *Annals of neurology* 20-21. Dans l'introduction, on vous dit qu'il y a à peu près 22 000 patients avec l'ataxie de Friedreich dans le monde, 5 000 aux États-Unis, et dans un rapport d'experts, il n'y en a que 825 en France. En quelque sorte, cet essai a été fait sur 11 centres, non pas en France, donc cela ne couvre pas la France, mais voici les chiffres. 22 000 patients total, tous les inclure demande un investissement absolument considérable de la part de l'industriel. En fait, il y en a 40 qui ont été inclus dans l'essai, pas plus. Nous sommes sur un sous-ensemble du sous-ensemble du sous-ensemble, ce qui vous donne les difficultés d'interpréter les données.

M^{me} CHEVREUL.- Tu voulais dire quelque chose.

Un chef de projet, pour la HAS.- Juste un complément, comme aujourd'hui, c'est votre accueil de bienvenue encore une fois et souvent, on a des collègues du médico-technique qui accompagnent les collègues. Par exemple, aujourd'hui, pour des raisons organisationnelles et des contraintes d'organisation, il n'y a pas un collègue ou les deux collègues qui ont supervisé le dossier à la CT, mais dans le futur aussi, vous bénéficiez de cette formule. Quand il y a un complément sur ce qui a été dit à la CT, on invite une personne senior, ou la personne qui a instrumenté le dossier. On l'a fait souvent, notamment pour les vaccins et pour les essais. C'est vrai qu'aujourd'hui, une présence d'une personne de la CT, un collègue aurait pu aussi vous donner plus d'éléments, plus synthétiques, puisque ces personnes ont assisté au vote. Dans le futur, on le fera.

M^{me} CHEVREUL.- Pour le coup, ça aurait pu arriver ou ça n'arrive pas, mais la discussion qui doit s'engager, c'est comment on le systématise et comment on fait qu'il y ait le minimum requis d'informations où on puisse aller le chercher. C'est déjà un premier champ d'avancée, de discussion pour la suite.

M^{me} CHAIX-COUTURIER.- J'ai une question de béotienne, je suis désolée. Je voudrais comprendre. On parle de patients au niveau monde. On parle d'une demande pour la France. Où est l'international ? Parce que les modélisations, ils l'ont fait pour la France. Ils la font au niveau monde. Je suis désolée, je ne suis pas au clair avec ça. Ce n'est pas du tout clair ce que je vous ai demandé ? Je n'arrive pas à comprendre l'intérêt qu'on a à un industriel à faire un truc absolument illisible.

M^{me} CHEVREUL.- Il n'a pas le choix.

M^{me} CHAIX-COUTURIER.- Mais pourquoi il n'a pas le choix ?

Une cheffe de projet, pour la HAS.- Parce qu'il est au-dessus de 20 million d'euros, donc il tombe sur les critères d'éligibilité.

M^{me} CHAIX-COUTURIER.- Non, mais par rapport, par exemple, à la population ciblée. On nous dit d'habitude, c'est fait comme ça, c'est super simple, ça devrait être comme ci, comme ça, machin, mais là il fait un truc...

Une cheffe de projet, pour la HAS.- Parce qu'on est dans le cadre d'une maladie rare, et qu'il a très peu de données. C'est pour ça que ce n'est pas facile de commencer par cet avis, parce qu'en fait, il n'y a rien de plus difficile que ces avis.

M^{me} CHAIX-COUTURIER.- Non, mais ça, j'ai bien compris.

Une cheffe de projet, pour la HAS.- Oui, mais d'où le peu de patients, le peu de données, du coup il bricole.

M^{me} CHAIX-COUTURIER.- On est d'accord,

Une cheffe de projet, pour la HAS.- C'est pas bien de dire ça, mais on aurait fait pareil ou peu ou prou, c'est ce que je dirais dans mon rapport. Il a fait ce qu'il a pu avec les données qu'il avait, il n'avait pas grand-chose, et encore, pour une maladie rare, il avait tout de même beaucoup de données. Ce n'est pas toujours le cas. Parfois, ils n'ont vraiment pas grand-chose. Il avait quelques données, mais malheureusement, là, beaucoup de critiques sont émises, mais ce n'est pas parce qu'il a fait un mauvais travail, c'est juste qu'il a fait avec ce qu'il a pu. Ce n'est pas facile comme dossier pour commencer et pour les chefs de projet et pour les membres de la CEESP.

M^{me} CHEVREUL.- C'est un vrai saut dans la commission, mais c'est aussi important de rappeler ce que ça veut dire une réserve. C'est-à-dire que la réserve, ce n'est pas forcément sur le fait qu'il est mal fait et c'est pour ça qu'il faut absolument qu'on le mette dans le rapport d'avis en disant qu'on le module. La réserve, c'est aussi sur comment ça peut ressembler à un vrai résultat, et comment ça peut ressembler à un vrai niveau d'efficacité par rapport à ce qui est un peu plus compliqué. C'est pour ça qu'on va attendre le rapport de Jennifer pour plus se mettre en perspective. On va laisser continuer la cheffe de projet.

Une cheffe de projet, pour la HAS.- Ensuite, je vous propose de passer aux paramètres d'estimation de l'efficacité. Il existait plusieurs sources de données, mais l'industriel a décidé d'avoir recours à une étude d'appariement par score de propension des données à trois ans de l'étude MOXle Extension et de la cohorte historique FA-COMS. À noter que dans cette étude, il y a une particularité. Tout d'abord, dans la population du bras omaveloxolone, on a une hétérogénéité entre l'antériorité au traitement des patients.

En effet, on a des patients qui étaient traités par omaveloxolone dans l'étude MOXle 2 et on a des patients qui n'étaient pas traités par omaveloxolone, qui étaient sous placebo dans l'étude

MOXle partie 2 et également dans l'étude MOXle partie 1. Ces patients n'ont pas la même antériorité au traitement.

Concernant la méthode de cette étude d'appariement. Le critère principal d'efficacité, c'est la variation du score mFARS à trois ans et les co-variables utilisées pour estimer les scores de proportion pour l'appariement sont le sexe, l'âge initial, l'âge d'apparition de la maladie, le score mFARS initial, le score de marche initial. Ces différentes co-variables ont été estimées au regard des données disponibles dans chaque étude, bien évidemment, et sur la base d'avis d'experts cliniciens selon l'industriel.

L'industriel justifie le choix de cette source de données en expliquant que, grâce à l'étude MOXle extension par rapport à l'étude MOXle partie 2, on dispose d'un recul plus important puisque l'étude d'extension va jusqu'à la 144e semaine versus 48e semaine pour l'étude MOXle partie 2 et que la méthode d'appariement est robuste.

L'industriel précise également que, étant donné l'hétérogénéité quant à l'antériorité au traitement, des analyses en sous-population ont été conduites selon cette antériorité au traitement et que les résultats sont globalement similaires. Notre analyse sur ce point, c'est qu'effectivement, l'hétérogénéité en termes d'antériorité au traitement est tout de même un point important et qu'aujourd'hui, il n'existe pas d'analyse statistique robuste qui permet de mettre en évidence que les résultats sont identiques dans les deux sous-populations. Par conséquent, il persiste une incertitude quant à l'impact de cette hétérogénéité dans le bras omaveloxolone.

De plus, concernant le choix des co-variables pour l'estimation des scores de propension, la variable du triplé GAA1 qui correspond selon l'industriel à la variable qui reflète le mieux la sévérité de la maladie n'a pas été incluse dans les co-variables pour estimer les scores de propension. Enfin, dans cette étude par appariement, il est fait l'hypothèse que la prise en charge et les méthodes de mesure du score mFARS sont identiques au cours du temps. En effet, on a tout de même dans la cohorte FA-COMS une cohorte qui est historique et qui a débuté en 2003, alors que l'étude MOXle d'extension a été réalisée entre 2018 et 2022.

Au regard de l'ensemble de ces éléments, il semble que les résultats d'efficacité estimés peuvent être considérés comme incertains et sujets à de nombreux biais. Et, élément important, aucune analyse de sensibilité avec les données de l'essai clinique MOXle Partie 2 n'a été conduite par l'industriel, il n'est donc pas possible de définir l'impact de ce choix de cette source de données sur les résultats de l'analyse.

Au regard de l'ensemble de ces éléments, nous proposons une réserve importante qui est la suivante : « Estimation de l'efficacité relative à partir de l'appariement par score de propension non robuste étant donné de nombreux biais et non explorée en analyse de sensibilité. »

Une intervenante.- La variable dont vous parlez était disponible ? Elle n'a pas été utilisée, mais elle était présente ?

Une cheffe de projet, pour la HAS.- Oui, effectivement, elle était disponible, elle n'a pas été utilisée.

Un intervenant.- Y compris dans la cohorte ?

Une cheffe de projet, pour la HAS.- Oui, je crois qu'elle était disponible dans la cohorte.

Concernant la mortalité dans le modèle, il est proposé de considérer deux types de mortalité : la mortalité spécifique liée à l'atteinte cardiaque, puisqu'il s'agit finalement de la principale cause de décès dans la pathologie, ainsi que la mortalité non spécifique. Les sources de données pour estimer la mortalité spécifique sont notamment l'étude Pousset et al., qui est l'étude rétrospective française et le PNDS.

En termes de méthodes d'estimation, tout d'abord, l'industriel propose d'estimer une probabilité d'être associé à un groupe qui va avoir une évolution rapide de l'atteinte cardiaque et, par la suite, d'estimer la survie de ce groupe de patients à partir de plusieurs hypothèses cliniques. Similairement à l'efficacité, on part du postulat de dire que les patients avec une atteinte cardiaque décèdent précocement, la plupart du temps autour d'environ quarante ans, ensuite que le risque d'évolution rapide de l'atteinte cardiaque est notamment dépendant du nombre de répétitions du triplé GAA1 qui reflète la sévérité de la maladie.

Par la suite, pour ces patients, il est proposé d'extrapoler les données de survie qui sont disponibles dans l'étude Pousset par une fonction Weibull. Effectivement, dans l'étude Pousset Pousset, on a une survie de l'ensemble des patients inclus dans cette étude, et on fait l'hypothèse que les décès observés avant 40 ans dans cette étude correspondent principalement à ces patients qui ont une atteinte cardiaque d'évolution rapide.

Le choix de la fonction Weibull par l'industriel est justifié rapidement en disant qu'il s'agit d'une fonction simple et qu'elle représenterait au mieux l'accélération de la dégradation de la fonction cardiaque observée dans l'étude Pousset. Par la suite, les paramètres de sa fonction sont encore une fois ajustés et calibrés pour correspondre finalement aux données observées dans l'étude Pousset, à savoir la survie moyenne et l'intervalle de confiance de cette survie observée dans cette étude.

Pour la mortalité non spécifique, on a quelque chose de relativement plus usuel avec le recours aux tables de mortalité par âge et par sexe de l'INSEE, avec toutefois une hypothèse selon laquelle le reste de la population dans la maladie a une espérance de vie similaire à celle de la population générale, donc par conséquent, on intègre la survie de la population générale pour ces patients.

Concernant notre analyse critique sur ces points, plusieurs points. Tout d'abord, la méthode d'estimation du triplé GAA1 n'est pas détaillée dans le rapport technique, mais dans le modèle, il est possible d'identifier comment ce point est estimé.

Enfin, le choix d'une fonction Weibull pour extrapoler les données de mortalité n'est pas justifié. En effet, on n'a aucune analyse de l'ajustement des données qui est présentée. On a seulement l'industriel qui nous dit que c'était la fonction la plus adaptée. Ce choix n'est pas exploré en analyse de sensibilité. On ne connaît donc pas l'impact sur le résultat, même s'il est attendu à faible.

Concernant la mortalité non spécifique, l'hypothèse selon laquelle la survie des patients sans atteinte cardiaque et du groupe de patients avec une atteinte cardiaque d'évolution stable serait identique à la population générale semble discutable. Étant donné l'espérance de vie attendue diminuée de ces patients à long terme, toutefois, étant donné l'horizon temporel qui est seulement de 20 ans, avec un âge moyen des patients de 26 ans, il n'est pas attendu d'impact important de cette hypothèse sur les résultats.

Au regard de ces éléments, nous proposons uniquement une réserve mineure sur « L'absence de justification du choix de la fonction d'extrapolation de la mortalité spécifique. »

M. SCHWARZINGER.- J'ai une réflexion méthodologique. Vous avez dit que dans le modèle principal, l'industriel ne prend pas en compte de modifications d'espérance de vie chez les patients avec le traitement. Si je fais l'analogie avec une autre maladie neurodégénérative beaucoup mieux connue, par exemple la maladie d'Alzheimer, ce qui fait qu'un traitement peut être « potentiellement » coût-efficace, c'est que vous retardez l'évolution dans la maladie parce que les patients meurent à date fixe. Là, c'est un peu ce qui est fait par analogie. Or, dans la diapositive d'avant, une corrélation est faite entre la cause principale de décès chez les patients et le score de sévérité, si j'ai bien compris.

Une cheffe de projet, pour la HAS.- Oui.

M. SCHWARZINGER.- Ça veut dire qu'il y a un biais.

Une cheffe de projet, pour la HAS.- Sauf que ce que je n'ai pas précisé, c'est que dans la population simulée, on simule des cohortes en monde parallèle. Nos patients ont exactement les mêmes risques d'avoir une sévérité de la maladie plus importante et on va modéliser l'évolution

du score mFARS. Mais typiquement, l'atteinte cardiaque qui est liée au nombre de répétitions de triplets GAA1 dans nos deux cohortes, traitement et meilleurs soins de support, on a les mêmes caractéristiques de patients.

Finalement, on a exactement les mêmes proportions de patients qui vont mourir, puisqu'on considère bien que le traitement n'a pas d'impact sur la survie des patients. Même si c'est lié à la sévérité, c'est modélisé de manière équivalente dans les deux bras.

M. SCHWARZINGER.- Pour des spécialistes, c'est crédible qu'un traitement puisse avoir un effet uniquement sur la partie neurologique, et pas du tout sur la partie cardiaque et la cause majeure de décès chez ces patients ?

Une cheffe de projet, pour la HAS.- Ce sont des hypothèses que nous avons discutées avec un expert clinicien. C'est vrai qu'on n'a peut-être pas poussé la question jusqu'à ce point. Nous avons vérifié les principales hypothèses qui étaient proposées par l'industriel. Il faut savoir aussi que nous sommes sur un horizon temporel de 20 ans. Le retour de l'expert clinicien était que la sévérité de la maladie était à long terme avec la survenue des manifestations cliniques autres que neurologiques. Je ne sais pas si cela répond à la question ou si tu veux compléter.

Une cheffe de projet, pour la HAS.- Ce qui nous a été rapporté, c'était que les patients qui allaient avoir une atteinte cardiaque et qui allaient décéder relativement jeunes, autour de 40 ans, c'était plutôt acceptable de considérer cela dans la modélisation. Et de considérer que l'autre partie de patients qui suivraient la mortalité de la population générale, en revanche cette hypothèse était peut-être un peu ambitieuse, puisqu'il y aurait de toute façon une espérance de vie un peu diminuée.

Ce qui n'est pas évident à différencier, c'est l'impact du traitement sur la mortalité ou le fait d'avoir un impact indirect du fait de dire : comme on ralentit, pense-t-on qu'il y aura un impact à terme sur la mortalité ou pas ? On s'est posé beaucoup de questions par rapport au type d'analyse au départ, mais vraiment aura-t-on un impact sur la mortalité ? Aujourd'hui, on ne peut pas répondre.

Une intervenante.- Du coup, c'est une hypothèse conservatrice, non ?

Une cheffe de projet, pour la HAS.- Oui, c'est ce que rapporte l'industriel, puisque comme aujourd'hui on ne sait pas s'il y aura un impact du traitement à long terme sur la maladie, le fait de ne pas le considérer, on demeure conservateur dans l'analyse, effectivement.

M. SCHWARZINGER.- Sauf que si vous rallongez l'espérance de vie, vous rallongez la durée de traitement, donc en termes de coûts, ce n'est pas vraiment...

Une cheffe de projet, pour la HAS.- On va arriver à la durée de traitement où on s'est posé beaucoup de questions.

M. SCHWARZINGER.- C'est plutôt exponentiel comme problème.

Une cheffe de projet, pour la HAS.- J'enchaîne, mais pas tout de suite sur la durée de traitement, mais une problématique qui est assez en lien, qui est ici la gestion de la dimension temporelle et notamment le maintien de l'effet traitement observé. Dans l'analyse, l'industriel propose de maintenir l'effet traitement, la réduction de la vitesse de progression du score mFARS sur tout l'horizon temporel. Ce choix est justifié par l'industriel puisqu'il dispose d'un recul de 144 semaines dans l'étude MOXle d'extension.

Notre analyse sur ce point, c'est que l'observation des données à 144 semaines ne permet pas de démontrer un maintien de l'effet traitement sur 20 ans, qu'il n'y a aucune discussion de la part de l'industriel sur la possibilité de mise en place de mécanismes d'échappement au traitement et qu'en l'absence d'analyse de sensibilité conduite de manière adéquate, l'impact de ce choix sur les résultats est inconnu.

Au regard de l'ensemble de ces éléments, nous proposons une réserve importante sur « L'application d'un maintien de l'effet traitement sur toute la durée de la simulation qui est incertain et qui est conjuguée à une absence d'analyse de sensibilité, explorant l'incertitude autour de ce choix. »

Maintenant, on passe aux événements intercurrents. Tout d'abord, concernant les événements indésirables, dans le modèle, il est proposé de modéliser uniquement les événements indésirables graves observés dans l'essai clinique MOXle partie 2 et uniquement pour le bras omaveloxolone. Ce choix semble acceptable dans la mesure où l'on a les meilleurs soins de support qui accompagnent l'omaveloxolone dans les deux bras de traitement.

Toutefois, concernant le choix de modéliser les événements indésirables graves, nous avons pu observer que certains des événements indésirables retenus avaient une sévérité légère ou modérée, et donc selon nous ne correspondaient pas forcément à des événements indésirables qui pourraient nécessiter une hospitalisation. Sur ce point, nous n'avons pas d'analyses de sensibilité qui sont conduites par l'industriel. Toutefois, on s'attend à un impact faible étant donné ce que le poste de coût de la prise en charge des événements indésirables peut représenter par rapport au coût du traitement, par exemple, ou aux autres postes de coût.

Au regard de l'ensemble de ces éléments, nous proposons la réserve suivante qui est « Absence d'analyse de sensibilité explorant le choix de retenir uniquement les événements indésirables graves comme proxy des événements indésirables nécessitant une prise en charge hospitalière. »

Ensuite, dans le modèle, il est fait le choix de modéliser uniquement les arrêts de traitement par omaveloxolone qui ont été observés jusqu'à trois ans dans l'essai clinique MOXle Extension. Ces arrêts de traitement correspondent à des arrêts de traitement en raison de survenues indésirables.

Notre analyse critique sur ce point, c'est que l'industriel ne discute pas la possibilité d'arrêt de traitement pour autre raison que la survenue d'événements indésirables. En effet, on pourrait envisager une possibilité de mécanisme d'échappement comme cela a été précédemment évoqué, ou un stade de la maladie trop avancé chez certains patients qui nécessiterait d'arrêter le traitement puisque les bénéfices apportés ne seraient plus suffisants.

L'industriel ne justifie pas également le choix de ne considérer des arrêts de traitement que jusqu'à trois ans. On n'a aucune analyse de sensibilité qui explore la possibilité d'un arrêt de traitement chez une proportion de patients au-delà de trois ans au fur et à mesure du temps. On a donc une durée de traitement qui est maintenue sur la totalité de l'horizon temporel au-delà de trois ans.

Au regard de ces éléments, il nous semble que « L'application d'une durée de traitement sur tout l'horizon temporel non explorée en analyse de sensibilité ayant un impact attendu important sur les résultats doit faire l'objet d'une réserve importante. »

Enfin, un dernier point, très rapidement, sur les manifestations cliniques concomitantes autres que neurologiques. Il est proposé dans le dossier de ne pas les considérer comme étant des événements intercurrents, car ces manifestations sont non résolutive. Ce point nous semble justifié et acceptable, toutefois, une limite à citer est que la modélisation de l'histoire naturelle de la maladie est limitée.

Une cheffe de projet, pour la HAS. - On vous propose de continuer ensuite sur la qualité de vie. Tout d'abord, dans l'essai clinique MOXle, il n'y a pas de questionnaire EQ-5D qui ait été administrée dans l'essai clinique, mais un autre questionnaire générique, le SF-36, qui a été donné aux patients. À partir de cela, n'ayant pas de questionnaire EQ-5D dans l'essai clinique, c'est vrai qu'en général, on privilégie une approche par mapping sous réserve qu'il existe une matrice méthodologiquement robuste et valide.

Ce qui est proposé dans le dossier, c'est que cette approche par mapping n'est pas privilégiée en analyse de référence par l'industriel, mais c'est plutôt un recours à la littérature qui est proposé en analyse de référence, alors que généralement, on propose plutôt la littérature en dernier recours et on privilégie un mapping s'il est validé. On a eu un peu de débats sur cette question.

La deuxième chose sur la prise en compte de la qualité de vie dans le modèle, c'est que, comme il vous a été présenté tout à l'heure par rapport à la structure du modèle, le fait qu'on ait une évolution continue des patients, la même chose est retenue pour la qualité de vie. Il y a une hypothèse d'une évolution de la qualité de vie qui est continue et linéaire et qui est corrélée à l'évolution du score SARA avec la corrélation SARA-score mFARS.

Sur cette hypothèse qui permet de prendre en compte la qualité de vie, on peut discuter sur le fait de considérer que ce soit linéaire et continu pour deux raisons. Si on regarde d'abord le côté statistique, dans la relation qui nous est présentée dans l'analyse de la littérature, dans l'étude de Reetz et al., on n'a pas de démonstration statistique sur cette linéarité. D'un point de vue clinique, après discussion avec l'expert clinicien, la qualité de vie est extrêmement importante à prendre en compte pour ces patients. Elle est très dépendante de la perte des capacités fonctionnelles et surtout très variables suivant les patients. Avoir une évolution linéaire considérée pour les patients ne serait pas forcément adapté ici.

Sur ces deux points, on a proposé une réserve, une sur le fait de considérer l'évolution linéaire et continue de la qualité de vie, alors qu'a priori, d'un point de vue de plausibilité clinique, ce n'est pas forcément adapté. Et sur le recours des données sur la littérature, plutôt que d'estimer les scores d'utilité à partir du mapping du questionnaire de l'essai clinique SF-36 vers l'EQ-5D. Néanmoins, l'analyse de sensibilité considérant ce mapping a été fournie par l'industriel et montre une augmentation des résultats.

Le recours au mapping, c'est-à-dire de prendre le questionnaire SF-36 de l'essai clinique, de le transformer en EQ-5D par un algorithme issu de la littérature conduit à augmenter les résultats de 72 %. Sur le graphique en bas à droite, vous pouvez voir la différence d'estimation du score d'utilité suivant la méthode, c'est-à-dire la littérature ou le mapping.

L'expression continue de la qualité de vie est rapportée ici en bas par l'équation qui relie le score EQ-5D au score SARA et au score mFARS dans le modèle.

M. DUPORT.- Il existe des échelles plus spécifiques pour les ataxies de Friedreich sur l'évaluation de la qualité de vie. Avez-vous demandé à l'industriel s'ils avaient utilisé ce type d'échelle ? Je pense au Friedreich's Ataxie Health Index, par exemple.

Une cheffe de projet, pour la HAS.- Le souci est que dans la modélisation, on a vraiment besoin de l'EQ-5D plutôt qu'une échelle spécifique pour pouvoir introduire un score comme dans les autres avis économiques et d'avoir à chaque fois l'EQ-5D. Pour l'échelle spécifique qu'on n'a pas spécialement demandée.

M. DUPORT.- Ce que je veux dire, c'est que je sais bien qu'on utilise l'EQ-5D, et que c'est beaucoup plus simple. C'est un débat qu'on a de façon éternelle ici, mais la généricité de l'EQ-5D fait que cela reflète très peu la situation de l'ataxie de Friedreich où les patients se plaignent de fatigabilité qui n'est pas évaluée par exemple dans l'EQ-5D, qui se plaignent de troubles de l'élocution, et c'est quelque chose qui a une forte demande. Par exemple, le score mFARS, c'est un score sur cinq points parmi plus de 90 points d'élocution. Je voulais savoir s'il y avait eu justement l'exploration à titre indicatif pour voir un peu ce qu'il en était sur des choses qui correspondent vraiment à des attentes des patients.

Une cheffe de projet, pour la HAS.- Le score SARA est pris en compte ici dans l'estimation de l'EQ-5D avec leur équation. L'EQ-5D est égal à -0,017 fois le score SARA plus la constante. Mais effectivement, on n'a pas d'autres questionnaires ici qui sont mobilisés dans l'analyse.

Les désutilités liées aux événements désirables graves sont prises en compte par des désutilités issues de la littérature. Les désutilités liées aux manifestations cliniques ne sont pas prises en compte ici pour éviter un double comptage, en disant que l'évolution de la qualité de vie qui est observée dans la littérature prend finalement déjà en compte indirectement ces manifestations cliniques.

Une intervenante.- Je peux vous poser une question juste pour être bien claire. Vous êtes en train de nous dire qu'ils ont choisi pour leur cas de base une modélisation de la qualité de vie à partir de la littérature, mais on a tout de même aussi le mapping disponible dans l'analyse de sensibilité.

Une cheffe de projet, pour la HAS.- Oui.

L'intervenante précédente.- Donc c'est juste de dire lequel des deux est....

Une cheffe de projet, pour la HAS.- Exactement. C'est le débat qu'on a eu en disant : aurait-on voulu en référence privilégier la littérature ou un mapping ? En tout cas, les deux sont disponibles et l'analyse est fournie.

L'intervenante précédente.- D'accord. Très bien. L'autre question que j'ai c'est que la qualité de vie qui est faite à partir du score SARA, c'est une troisième analyse ?

Une cheffe de projet, pour la HAS.- Non. La littérature étant retenue en référence, c'est l'étude Reetz et al., qui est la cohorte européenne dans laquelle des questionnaires de qualité de vie ont été collectés.

L'intervenante précédente.- Ça revient à la question que vous posiez. C'est tout de même qu'ils attachent de l'importance à un indicateur différent de l'EQ-5D.

Une cheffe de projet, pour la HAS.- C'est en lien avec l'intégration des données ici. Dans l'intégration des données, on considère que la qualité de vie est liée et a également une évolution continue étant donné le modèle de microsimulation continue. Du coup, on rattache nos données d'EQ-5D dans cette équation pour considérer que lorsque notre score mFARS évolue, notre qualité de vie évolue également de manière continue. C'est plus de l'intégration que liée à la source donnée en tant que telle, puisqu'on aurait pu utiliser une source de données...

L'intervenante précédente.- Vous avez parlé du score mFARS. Ça veut dire qu'ils passent du score mFARS à un score SARA qui lui-même est ensuite utilisé pour faire l'EQ-5D.

Une cheffe de projet, pour la HAS.- C'est ça. Et pour le coup, la corrélation mFARS SARA est publiée et acceptée.

L'intervenante précédente.- C'est qu'ils attachent de l'importance à la qualité de vie mesurée par autre chose qu'un EQ-5D. Merci.

M. SCHWARZINGER.- C'est surtout parce que de facto, ça revient à la première question d'Antoine tout à l'heure sur l'effet clinique minimum. Vous avez forcément des effets plus marqués avec une échelle spécifique. C'est un éternel débat.

L'intervenante précédente.- Mais j'entends leur choix.

M. SCHWARZINGER.- Le guide méthodologique, j'imagine, c'est de dire qu'on choisit l'échelle générale EQ-5D, quelle que soit la maladie.

Une cheffe de projet, pour la HAS.- Oui, et qu'on privilégie les sources de l'essai clinique en premier lieu, et si jamais on ne peut pas, plutôt la littérature.

Une intervenante.- Ce qui n'empêche pas, si on a des résultats très divergents sur une échelle spécifique, de le mentionner pour exprimer le fait qu'on peut avoir un manque de captation par l'EQ-5D de l'effet sur la qualité de vie, si possible effectivement.

Une cheffe de projet, pour la HAS.- L'argumentaire aussi du laboratoire avec le SF-36, c'est qu'ils expliquent aussi que, finalement, il ne serait pas très cohérent dans les résultats qu'ils obtenaient, etc. Ils développent aussi un argumentaire pour dire pourquoi ils n'utilisent pas ce mapping.

Un chef de projet, pour la HAS.- C'est une bonne chose que vous discutiez ces éléments parce qu'on reviendra au fur et à mesure des réunions de la CEESP et des agendas pour parler des maladies rares. L'EQ-5D avec les maladies rares, c'est vraiment un chat fouetté.

Nos recommandations, c'est qu'on recommande une échelle générique parce qu'on a une matrice de pondération française pour l'utiliser pour les états de santé. Dans l'évaluation économique, on ne s'intéresse pas à la significativité de la qualité de vie. Même si c'est corrélé,

on s'intéresse plutôt aux états de santé. Dans un premier temps, a-t-on des échelles génériques de l'EQ-5D qui sont recommandées en première intention ? Le cas échéant, si on n'a pas l'EQ-5D, a-t-on une autre échelle générique, par exemple pour la population pédiatrique ou autre ? Puis, si on n'a pas une autre échelle générique, a-t-on une échelle spécifique qui a un pont, un *bridge* pour le transformer en échelle générique pour pouvoir utiliser la pondération des préférences françaises, ce qu'on appelle les tarifs français des index français ? Après, si on n'a pas un mapping, faute de mieux, le cas échéant, on patage un peu partout dans la littérature, que ce soit au NICE ou ailleurs. On dit que pour des maladies rares, on ne peut pas faire, on choisit dans la littérature, mais on regarde par rapport aux états de santé.

En fait, il y a un schéma séquentiel que nous suivons. D'ailleurs, NICE le suit aussi. Les Allemands ne font pas d'évaluation économique, les autres, donc on part toujours de l'échelle générique. C'est quelque chose qui fait que cette CEESP, dans son programme, dans le futur, aura des discussions pour alimenter dans le futur la discussion sur la qualité de vie pour les maladies rares. Souvent, on sait d'avance que l'EQ-5D, quoi qu'il en soit, quelle que soit la situation pour les maladies rares, pose toujours des soucis. Mais le problème, c'est qu'on est obligé de l'utiliser pour pouvoir utiliser les tarifs.

La question qui se pose dans le futur, c'est comment on peut, au-delà du quantitatif qu'on donne et qui est parfois un peu rébarbatif, parfois même un peu difficile, pour des gens qui ne sont pas spécialistes, comment on peut ajouter ce qualitatif. C'est quelque chose que nous espérons ajouter dans les futures recommandations du guide, comme pour le NICE, quand on dit les maladies rares, on dit il faut ajouter, comme disait la cheffe de projet, s'il y a d'autres dimensions qui ne sont pas renseignées.

C'est vrai que l'ajout qualitatif d'informations sur la qualité de vie dans les états de santé s'impose et est toujours discutable pour les maladies rares. On a des échelles spécifiques. Le problème, toutes les échelles spécifiques n'ont pas de mapping, et même parfois, s'ils ont du mapping, il faut voir le site d'Oxford pour les mappings qui sont validés par Wenn, par Brazzy, par d'autres.

Du coup, on a toujours un peu cette difficulté dans l'évaluation économique, c'est qu'on a besoin du quantitatif. Mais pour les maladies rares, c'est vrai, c'est une discussion qui va toujours se poser. Même pour cette CEESP, par la suite, dans nos doctrines, utilise-t-on comme résultat de santé le QALY pour les maladies rares ? Peut-on faire autre chose ? Il y a des belles discussions qui se profilent dans le futur pour ces choses.

Mais on a nos guides, et on ne sait jamais dans le futur, si on a des choses pour vous montrer comment on utilise les échelles, les questions qui se posent, nos problèmes aujourd'hui actuels par rapport au NICE. Parce que notre matrice de pondération française a été heureusement faite

par Gérard de Pouvreville et les gens qui ont financé, qui sont de l'industrie. Dans les autres pays, la matrice de pondération nationale a été financée par le public. Il y a plein de choses qui se posent pour la qualité de vie de manière générale où on a tout de même l'algorithme, mais pour les maladies rares, c'est vrai qu'on reste toujours sur sa faim. Je voulais vous donner ce petit aperçu. Karine, merci.

M^{me} CHEVREUL.- Merci, la question qui se pose aujourd'hui, c'est justement entre ces différentes options, comment on considère celles qu'ils ont choisies ? On attendra peut-être le premier retour de notre rapporteur pour en discuter plus avant. Hassan, je te laisse la parole.

M. SERRIER.- Juste une petite question de complément, c'est vrai que c'est une posture qui est un peu inhabituelle. Vous allez prendre l'habitude de le faire, mais la question de fond, c'est : ont-ils les arguments ? Parce que j'entends le fait que peut-être que la littérature finalement reflète mieux la réalité que le mapping à partir d'un SF-36, sauf qu'il faut les arguments pour le prouver. En l'absence d'arguments, il vaut mieux aller vers le conservateur, et le conservateur, c'était l'autre. Surtout c'est impressionnant de voir +71 %, et ma question portait là-dessus. +71 % du RDCR, mais le RDCR peut être très instable, a-t-on la variation des QALY liée à cela ? Sinon, pour plus tard, vous nous le donnez.

Une cheffe de projet, pour la HAS.- On a un différentiel de QALY de 0,36, alors qu'en référence 0,61. En référence on a 0,61 de différentiel et on a 0,36 avec le mapping.

M. SCHWARZINGER.- Juste pour être clair, 0,61, cela veut dire quoi ? C'est six mois en bonne santé, sur trois ans. C'est énorme.

Une cheffe de projet, pour la HAS.- Peut-être que sur la validation on vous propose d'aller vite parce qu'effectivement l'heure tourne. Sur la validation, on est dans un contexte d'une maladie rare que les données observationnelles ont déjà été utilisées pour la majorité dans la modélisation, ce qui fait que pour la validation externe, ce n'est pas évident de disposer encore de données dans la littérature pour venir valider les différentes hypothèses qui sont faites dans l'analyse.

Ensuite sur les coûts, l'évolution continue de la maladie est faite un peu différemment. Des catégories sont faites, des états de santé sont réalisés pour prendre en compte les coûts en définissant des seuils sur l'échelle SARA sur avis d'experts. Le souci c'est que l'on a une variabilité fonctionnelle importante, et que pour un même score mFARS, on peut avoir une certaine variabilité dans les atteintes cliniques des patients. C'est ce qui fait que l'on est face à une difficulté de faire des catégorisations de coûts avec ces états de santé, alors que finalement, un score peut conduire à plusieurs états cliniques.

Autrement, sur la prise en compte des manifestations cliniques, on a vu tout à l'heure que celles autres que neurologiques n'étaient pas prises en compte, notamment dans les désutilités, mais sont prises en compte pour les coûts. Là l'on a une certaine hétérogénéité, des fréquences qui sont un peu complexes à identifier, qui manquent un peu de précision dans ce qui est proposé.

On propose ici une réserve mineure sur l'estimation de ces fréquences des manifestations cliniques et une réserve importante sur l'approche catégorielle des coûts qui demeure incertaine au regard de la forte hétérogénéité individuelle de l'expression clinique en fonction du score mFARS.

Je suis désolée, je vais un peu vite pour essayer de rattraper un peu le retard.

Les résultats de l'analyse de l'efficience. Une fois la modélisation tournée, on obtient sur 20 ans un RDCR de 4,7 millions euros par QALY de l'omaveloxolone par rapport à la prise en charge par meilleurs soins de support. On a également les analyses de sensibilité sur le prix qui sont faites et qui varient dans le même sens que la variation de prix de l'omaveloxolone proposée. En termes de dispositions à payer, on a une probabilité de 80 % pour l'omaveloxolone d'être coût-efficace, pour une disposition à payer d'environ 6,4 millions d'euros.

Ces résultats, on vous propose de les reprendre un peu par la suite, puisqu'a priori, on vous proposerait de ne pas retenir ces résultats, en raison d'une incertitude trop forte, mais on y reviendra tout à l'heure. Des analyses de sensibilité déterministes ont été réalisées. On a le diagramme de Tornado qui est fourni et des analyses de sensibilité en scénario associé. Je suis désolée, on ne voit pas très bien sur la diapositive, mais plusieurs analyses ont été menées avec des variations qui tendent à faire diminuer le RDCR ou à augmenter le RDCR de manière significative.

Peut-être que l'on peut revenir par la suite si vous souhaitez sur ces différentes analyses proposées, et on vous propose de passer à l'impact budgétaire.

Une cheffe de projet, pour la HAS.- L'objectif de l'analyse d'impact budgétaire est d'estimer l'impact sur les dépenses de l'Assurance-maladie, de l'introduction sur le marché de omaveloxolone dans son indication. Cette analyse est conduite selon une perspective Assurance-maladie, sur un horizon temporel de trois ans. La population d'analyse est la population de l'indication de l'AMM, à savoir les patients adultes et adolescents âgés de 16 ans et plus. Les scénarios comparés sont un scénario sans SKYCLARYS qui retient uniquement les meilleurs soins de support et un scénario avec SKYCLARYS dans lesquels on retrouve SKYCLARYS et les meilleurs soins de support. L'ensemble de ces choix structurants sont considérés comme acceptables.

Concernant le modèle d'impact budgétaire, ici, il est proposé une approche de cohorte ouverte de patients prévalents. Chaque année, on a une population prévalente qui est intégrée et qui prend également en compte les arrêts de traitement. La structure d'un modèle s'apparente davantage, selon nous, en termes d'analyse critique, à un modèle de cohorte incidents prévalents. En effet, on intègre les patients qui sont nouvellement diagnostiqués et adressés chaque année vers les filières de maladies rares. C'est un point sur lequel on reviendra un peu dans l'estimation de la population cible pour vous décrire à quoi correspondent ces patients incidents.

Bien que ce choix est peu décrit, il est considéré comme acceptable.

Enfin, l'analyse d'impact budgétaire repose sur une modélisation similaire à celle de l'analyse d'efficience. Par conséquent, on retrouve des réserves en miroir de l'analyse d'efficience. Notamment sur les données cliniques, on ne retrouve pas la totalité des réserves importantes et mineures, mais seulement celles qui sont jugées à avoir un impact sur un horizon temporel de trois ans. Sur les données de coûts, on retrouve également les deux réserves, une importante et une mineure, présentées précédemment par la cheffe de projet.

Une cheffe de projet, pour la HAS.- Pour estimer la population cible dans l'analyse d'impact budgétaire, tout d'abord, on part de la prévalence de l'ataxie de Friedreich identifiée dans le PNDS en Europe, appliquée au nombre d'habitants en France de plus de 16 ans. Cela donne une population estimée environ à 1 850 patients. Les données de la base nationale des maladies rares ont été mobilisées, qui rapportent 705 patients de plus de 16 ans ayant un diagnostic confirmé au probable d'ataxie de Friedreich et qui sont adressés et diagnostiqués dans les deux filières maladies rares, FILNEMUS et BRAIN-TEAM.

Ensuite, une hypothèse de l'augmentation de la population cible recensée dans cette base nationale des maladies rares dans le contexte de la mise à disposition de l'omaveloxolone, ferait que le fait d'avoir un traitement disponible augmenterait le nombre de patients qui pourraient être traités. Cela conduit à une estimation de la population diagnostiquée et prise en charge en année 1 de ■■■ patients, ■■■ patients en année 2 et ■■■ en année 3.

C'est l'hypothèse du laboratoire. Il n'y a pas plus de sources de données autres qu'une hypothèse.

Pour les parts de marché, dans le scénario sans omaveloxolone, 100 % des patients sont traités par meilleurs soins de support, et dans le monde, avec omaveloxolone plus meilleurs soins de support, l'industriel estime une part de marché à ■■ % en année 1, ■■ % en année 2 et ■■ % en année 3.

De manière assez classique, c'est vrai que ces parts de marché restent arbitraires. En revanche, on s'attend à un taux de pénétration sur le marché assez rapide et important étant donné d'une absence de traitement disponible aujourd'hui. Les parts de marché sont variées de plus ou moins 10 % et plus ou moins 15 %.

En termes de résultat de cette analyse d'impact budgétaire, on a une population rejointe en année 3 de ■■■ patients et un impact budgétaire estimé à 544 millions d'euros sur un horizon temporel de trois ans. Si on regarde de manière plus précise les différents postes de coûts de l'impact budgétaire, le coût de traitement représente à peu près ■■■ % de l'impact budgétaire total.

On a des analyses de sensibilité déterministes proposées et des analyses de sensibilité en scénario et également sur le prix. L'impact budgétaire va varier proportionnellement avec le prix, la taille de la population cible prévalente diagnostiquée et prise en charge, et les parts de marché.

Je ne sais pas s'il y a encore des questions. Désolée, on est allé un peu vite sur la fin.

M. SERRIER.- Oui, j'ai juste une question de compréhension. Je suis peut-être passé à côté de l'information, mais vous avez parlé du traitement ■■■ euros, 90 gélules, ça représente combien par an, combien ils prennent de gélules, et l'idée du coût annuel ?

Une cheffe de projet, pour la HAS.- De tête, je ne sais pas. On va pouvoir trouver.

M. SERRIER.- 90 gélules par mois ?

Une cheffe de projet, pour la HAS.- Ce sont des boîtes de 90 gélules.

Une cheffe de projet, pour la HAS.- C'est 3 gélules par jour, donc c'est bien 90 gélules par mois pour un mois.

M. SERRIER.- ■■■ euros pour un mois ?

Une cheffe de projet, pour la HAS.- Oui.

M. BENARD.- Il n'y a pas de prise en compte des événements indésirables, il y a un profil de sécurité qui est très bien ?

Une cheffe de projet, pour la HAS.- Les événements indésirables sont pris en compte dans l'analyse, dans les coûts et dans les utilités liés aux événements désirables graves et aux manifestations cliniques non. Mais oui, les événements indésirables sont considérés.

Une cheffe de projet, pour la HAS.- En résumé, si on reprend l'analyse de l'efficience, l'analyse critique conduit à avoir huit réserves importantes et cinq mineures. Pour l'analyse d'impact budgétaire, il y a trois réserves importantes et trois mineures.

S'il n'y a pas d'autres questions, nous vous proposons une proposition de conclusion, mais qui est effectivement soumise à la discussion et à débat. Ce qui est proposé aujourd'hui dans le projet d'avis, c'est de ne pas retenir les résultats de l'analyse de l'efficience en raison d'une incertitude globale majeure, si on reprend la doctrine, par rapport à un cumul de réserves importantes qui conduirait à des résultats non interprétables, et en détaillant les différentes sources d'incertitude du dossier.

En revanche, l'impact budgétaire, il est proposé de ne pas invalider les résultats, puisqu'un certain nombre d'hypothèses n'impactent pas de manière significative les résultats de l'impact budgétaire contenu de l'horizon temporel de trois ans. Il est proposé en revanche de présenter les résultats de l'impact budgétaire.

C'est soumis à discussion et à débat. N'hésitez pas.

M^{me} CHEVREUL.- Antoine.

M. BENARD.- Merci beaucoup c'était très instructif. Je ne l'ai pas dit tout à l'heure, mais c'est la première fois que j'assiste à la CEESP. La question n'est pas de remettre en cause le fait qu'il y ait une incertitude globale majeure : ça, c'est sûr. Je voudrais tout de même souligner l'effort de l'industriel qui a produit des résultats avec le peu de données qu'on a, et qui ne se cache pas d'afficher un ratio coût-utilité à 4 millions d'euros. C'est plutôt bien de sa part.

Mais je reviens sur ce ratio coût-utilité à 4 millions d'euros avec une probabilité d'efficience à zéro à un seuil de 2 millions d'euros par inaudible 3.24.11. Pourquoi on n'est pas plus direct dans les conclusions, en disant que ce n'est pas efficient ?

M^{me} CHEVREUL.- Pour le coup, je vais laisser la parole à Hassan maintenant, puis je vais te répondre, puisqu'Hassan fait comme ça. Je le laisse répondre, puis je réponds après.

M. SERRIER.- Je secoue le bras, parce que j'allais intervenir là-dessus aussi. C'est un point important, et ça relève d'une zone grise. Pour remettre dans le contexte de la CEESP, notre rôle, c'est de dire si on croit ou pas, je simplifie, dans l'évaluation qui est fournie par l'industriel. Peut-elle être utilisée par le CEPS ou pas ? Quand il y a un problème méthodologique majeur, une réserve majeure, on dit non, on n'y croit pas du tout, il y a un problème, on n'y croit pas du tout, donc laissez tomber. Souvent, ça arrive avec des ratios qui sont bas. On dit non, n'y croyez pas, ce n'est pas efficient.

Après, il y a le cas de figure où on a plein de petits problèmes, plein de choses, et finalement, on ne croit pas au résultat parce qu'on ne sait pas, il y a trop de choses, ce qui est le cas de ce dossier.

Malgré tout, et je suis toujours perturbé, parce que le fait de mettre une incertitude globale majeure, ça veut dire au CEPS : ne regardez pas le résultat. Pourtant, il y a beau avoir une incertitude globale majeure, on sait que ce produit n'est pas efficient.

M^{me} CHEVREUL.- Je ne sais pas si ça veut dire : ne regardez pas les résultats. Ça veut dire : faites-en ce que vous voudrez, c'est-à-dire certaines parties.

M. SERRIER.- En théorie, ça veut dire qu'on invalide le résultat.

M^{me} CHEVREUL.- Ce n'est pas la même chose, on invalide le résultat, mais ce sera très intéressant qu'on en reparle avec eux, et c'est aussi une des occasions de la CEESP de le faire, que de dire il y a une incertitude globale majeure, ça ne vous empêche pas de le regarder morceau par morceau.

M. SERRIER.- Ce qui avait été discuté, c'était : ce n'est pas parce que vous invalidez le résultat global qu'on ne va pas regarder des choses intermédiaires, par exemple. Normalement, ils ne peuvent pas utiliser le résultat si on l'invalide. C'est pour ça que je suis toujours gêné quand c'est un ratio aussi élevé de l'invalider, parce qu'on sait que ce n'est pas efficient.

M^{me} CHEVREUL.- Pour le coup, doit-on l'invalider seulement sur la méthode où on pense qu'effectivement, il y a une incertitude sur le résultat ? Il y a une incertitude globale majeure parce qu'il est haut ou parce qu'il est bas, c'est déjà une deuxième prise de position de la CEESP. C'est-à-dire que la CEESP rend un avis méthodologique en disant : oui, on pense que le résultat est plutôt robuste, c'est un peu ce qu'on est en train de dire, ou il y a une incertitude qui est trop importante.

L'autre partie, et c'est celle que tu abordes, qu'on ne fait pas et qui est dans la doctrine, on peut après aussi rendre un avis sur le fait qu'on pense que ce n'est pas efficient ou c'est trop élevé ou autre. Or, ce n'est pas fait et c'est aussi un des sujets, ça me permettra, quand on aura fini, qu'on puisse parler des groupes de travail et ce qu'on peut faire après pour avancer sur ce point.

Celui-ci, de façon générale, on ne le fait pas parce que, même si dans la doctrine on dit on pourrait le faire, il n'y a pas de règle. Il n'y a pas de règle de raisonnement qui a été mise en place par la CEESP, pour pouvoir aller plus loin et dire : on trouve que c'est efficient ou pas. On n'est pas comme les Anglais puisque c'était le premier exemple avec le seuil, c'est au-dessus de temps, hop, ce n'est pas efficient et on ne discute pas.

Ce sera intéressant surtout pour nos amis qui sont plus novices ici dans la commission sur l'évaluation médico-économique, qu'on reparte de : comment on dit que c'est efficient ou pas et qu'on se dise quelle est la question qui se pose, mais c'est dans un deuxième temps. Là, pour le coup, tu te dis 4 millions, j'aurais tendance à dire que ça ne l'est pas, mais c'est parce qu'on est déjà formaté. La première réponse dans l'avis, c'est, techniquement, avec ce qu'ils ont fait, les

choix qu'ils ont faits, que ce soit pour la population, pour la modélisation, pour l'efficacité, pour l'évaluation des coûts et les différentes diapositives qui nous ont été présentées, thématique par thématique, avec la réserve technique qu'on a, croit-on à ce résultat de 4 millions ? Parce que si ça se trouve, c'est 8 ou si ça se trouve, c'est 1 ou même 500 000. J'exagère, mais ce qui n'empêche qu'effectivement, on émet notre avis là-dessus.

Puisqu'on a un cumul de réserves importantes et que vous nous proposez, parce que c'est une proposition qui est faite, une incertitude globale, se dit-on qu'avec toutes ces réserves, effectivement, le résultat est tout de même incertain ? Parce qu'il y a une trop grande variabilité dans l'échelle entre quoi et quoi, on ne sait pas. Parce que même l'analyse de sensibilité, on dit qu'ils prennent des choses, mais est-on vraiment dans cette fourchette ?

Doit-on ne pas dire qu'on a cette incertitude globale, parce qu'on se dit tout de même que quand c'est aussi élevé, ce serait dommage qu'ils ne le regardent pas ? Il faut absolument qu'on arrive nous-mêmes à avoir une façon de penser. C'est un, je regarde techniquement, j'y crois ou pas ? Deux, vais-je croire que c'est efficient ou pas ? Mais ce n'est pas parce que c'est très élevé qu'on se dit que ce serait dommage qu'ils ne le regardent pas tout de même.

M. SERRIER. - C'est juste qu'on a le sentiment qu'on a une information. On a eu par le passé aussi des cas de figure où on avait beaucoup de réserves comme ça. Par contre, et je crois que ce n'est pas le cas, vous me dites que c'est le contraire, tous les problèmes allaient dans le même sens, c'est-à-dire que toutes faisaient monter le RDCR. Donc, même si l'incertitude est très importante, on sait qu'à minima, le RDCR est de tant, et là, dans ce cas, ça ne pose pas de problème. Mais là, je crois que ce n'est pas le cas, je crois que ça va dans tous les sens.

Une cheffe de projet, pour la HAS. - Il y a vraiment des hypothèses où, pour le coup, la réponse, c'est : on ne sait pas. On est vraiment dans le flou en disant que durée de traitement, on n'en sait rien, finalement. Après, dans leurs analyses de scénarios, il faut être honnête, il y a aussi des analyses qui font baisser le RDCR. Toutes ne font pas augmenter le RDCR, mais c'est vraiment suivant le choix qui est fait dans la modélisation.

M^{me} CHEVREUL. - Alexandre.

M. BARNA. - Si on a le temps et si vous me permettez que je mette les pieds dans le plat encore plus. À l'occasion de ce dossier, je me posais la question en termes de cohérence entre la CEESP et la CT, carrément s'il y a des QALY ou si c'est zéro QALY. Comment l'expliciter ? On a appris que la CT s'est prononcée pour une ASMR de niveau V, c'est-à-dire pas d'amélioration du service rendu, c'est-à-dire delta bénéfice, delta conséquences évalué encore une fois par la CT, je sais très bien ce que c'est, égal zéro. C'est en plus renforcé par un message sur le SMR, si j'ai bien

compris, qui va être modéré, c'est-à-dire tout de même enfoncer le clou. On voit très bien les médicaments qui ont un SMR modéré.

Encore une fois, je sais bien ce que c'est dans les ASMR. De l'autre côté, si on regarde les QALY, je connais un peu ce qu'il y a dans les QALY, c'est une multiplication de la qualité de vie par la quantité de vie, c'est là où je veux en venir. Les gens vont sauter de leur chaise en disant qu'il y a une différence, mais justement, je sais qu'il y a une différence. Il n'empêche, ce sont deux indicateurs qui mesurent d'une certaine manière un peu la même chose, c'est-à-dire le bénéfice, les conséquences en termes d'efficacité, que ce soit *efficacy*, *effectiveness*, là aussi ce n'est pas le moment, en termes de conséquences de ce médicament sur l'amélioration de l'état de santé.

C'est là où je veux en venir. C'est vrai que d'une certaine manière, cela me gêne, mais c'est peut-être un sujet à débattre, ce n'est pas pour tout de suite, que d'un côté, la CT se prononce pour une ASMR de niveau V, c'est-à-dire delta égale zéro par rapport à l'existant, et de l'autre côté, nous, même si on arrive à un coût par QALY de 4,5 millions, n'empêche que le dénominateur qui est QALY est, par définition, positif. À un moment donné, peut-être réfléchir en termes de cohérence entre l'ASMR et le QALY.

M^{me} CHEVREUL.- Je vais laisser Alexandre répondre, mais ce n'est pas tout à fait l'objet aujourd'hui, même si c'est une vraie réflexion qu'il faut avoir et que l'on va engager. Mais c'est aussi pour cela qu'il y a deux commissions séparées, et c'est aussi pour cela qu'on a toutes ces questions sur la signification clinique de l'amélioration que l'on a vue, et c'est aussi pour cela que c'est tellement petit que l'on arrive à 4 millions, ou dans la marge des 4 millions, c'est bien le reflet.

Après, on peut se poser la question, et c'est pour cela que ce serait intéressant qu'on ait une rencontre avec le CEPS pour qu'il vous reparle de comment ils utilisent cela. C'est-à-dire que, quand vous avez une incertitude globale, le résultat ne peut pas lui être opposé, par l'industriel. Cela ne veut pas dire qu'il n'a pas le droit de le regarder s'il a envie de le regarder, parce qu'on va lui donner tout de même. Ce qu'il va en faire, je n'en sais rien, et je ne sais pas s'il y aurait une différence. Je te laisse la parole, Alexandre.

M. BEAUFILS, pour la HAS.- Oui j'allais prolonger. Il faut faire attention, les commissions sont tout de même indépendantes, il faut faire aussi attention à se dire « puisque la CT a dit comme ci, il faudrait adapter et mettre comme cela ». Non, mais je caricature aussi, ce n'est pas ce que vous avez dit. Il faudrait faire attention parce que la CT a sa propre doctrine d'évaluation, la CT n'a pas donné un SMRI, elle a donné un SMR modéré, elle a tout de même dit « OK » pour le remboursement, elle n'a pas dit « il y a zéro ». Après, de dire : c'est faible, c'est important, c'est modéré ou c'est SMRI, on a le même dossier sous les yeux, et on le regarde de manière différente.

Aujourd'hui, et Karine l'a dit, on a les mêmes questions que la CT, puisque la modélisation qui est faite à partir de l'efficacité, la qualité de vie et les coûts fait que l'on a un RDCR sous les yeux de 4,7 millions. Cela veut bien dire que l'on a finalement les mêmes questions sous-jacentes si je caricature des coûts importants pour un différentiel d'efficacité assez faible, je caricature peut-être à l'excès. Je prends des pincettes, mais c'est ce qui donne d'un côté une évaluation qui est effectivement peut-être jugée faible, mais qui est quand il y a une question positive de la CT, et qui donne de notre côté, avec un prisme différent, en ajoutant de la qualité de vie, un RDCR à 4,7 millions, et toutes les questions qui sont : que dit-on de ce RDCR à 4,7 millions ? On a le même dossier sous les yeux, et on est indépendant, on voit bien les mêmes choses, le même problème.

M. BARNIA. - Ce n'est pas le moment de discuter de ça.

M^{me} CHEVREUL. - Oui, par contre, si on est tous là, c'est pour pouvoir générer une discussion qui soit plus une discussion dans un groupe de travail pour qu'on en reparle. C'est aussi pour cela que vous êtes tous là et qu'on est tous là. Par contre, il est important de se dire que, globalement, même si ce RDCR est haut, les résultats qui nous sont proposés, vu les incertitudes qu'il y a, les potentielles inexactitudes sur certains choix, il serait bien de se dire que c'est vraiment 4 millions, plutôt que de se dire que c'est plutôt 1,5 million ou 8 millions ? C'est ça le sujet.

Là, on peut dire effectivement, mais je vais laisser enfin laisser Jennifer parler, puisque c'est elle qui est rapporteure, il y a tellement d'incertitudes qu'on ne peut pas dire que c'est 4 millions, l'incertitude est très importante. C'est ce que vous vouliez dire, n'est-ce pas ? Sandy, je vais te laisser intervenir, après, on va laisser Jennifer tenir son rôle de rapporteure. D'accord.

M^{me} TUBEUF. - Excuse-moi Jennifer. C'est intéressant de parler d'incertitude, néanmoins, par exemple, ici, ce n'est pas nécessairement négligeable, sept mois de vie supplémentaires, surtout en bonne santé, pour des personnes qui ont déjà une expérience de vie qui peut être limitée. Selon moi, ça peut être intéressant. Mais quand on parle d'incertitude, j'aime bien aussi avoir des intervalles de confiance. Comme tu disais tout à l'heure aussi, il y a des choses qu'on aurait voulu et qu'on n'a pas ici, j'aurais bien aimé savoir, justement dans cette estimation des QALY, l'intervalle de confiance autour de cet estimateur. Ça peut être présenté sous la forme d'un nuage de points dans un plan de coût-efficacité, mais ce serait bien d'avoir ces éléments. Parce que quand on veut l'incertitude, on veut savoir aussi quelle est l'ampleur des différences qui s'observent, et on ne les a pas ici.

M^{me} CHEVREUL. - Je vous propose qu'on continue après que Jennifer a eu la parole, si vous êtes d'accord. Jennifer, vous êtes d'accord ?

M^{me} MARGIER. - Je suis tout à fait d'accord, mais je n'ai pas tellement plus de choses à dire. Je voulais juste rapporter et remercier les chefs de projet. Vous avez constaté que c'était un énorme

travail. Pour vous refaire l'historique, en plus, ça n'a pas été facile pour elles, parce qu'on a commencé il y a très longtemps, et il y a eu une pause dans ce dossier. C'est très dur de revenir dans un dossier. On l'a bien en tête quand on vient de le lire et quand on échange avec les industriels, mais après, le temps passe et c'est tout de même pointu. Cela n'a pas été évident pour elles. Elles l'ont bien fait, elles ont bien rapporté, parce que vous avez posé plein de questions, vous aviez au moins la base pour comprendre le minimum de ce modèle. Je les remercie pour ça. Elles ont fait un travail de grande qualité.

Je voulais juste revenir, pas sur les aspects techniques, on en a déjà vu beaucoup, mais plus sur l'idée globale du projet. Il y a eu huit réserves importantes, c'est beaucoup. Il n'y en a pas toujours autant. Il n'y a pas de jugement de valeur quand on rapporte ou quand on évalue le dossier, c'est conforme, ce n'est pas conforme, ça génère de l'incertitude, oui, non. On n'est pas en train d'attribuer des bons points aux industriels à dire : ils ont bien fait leur travail, là non. Ce n'est pas l'idée. On est complètement impartial. Ce n'est pas la question.

Ça ne veut pas dire que parfois, on se dit tout de même : ils ont un peu abusé. Ça arrive qu'on se le dise tout de même entre nous, plus d'une fois. Mais globalement, on n'est pas là pour ça. On essaie d'être le plus impartial possible. Justement, je suis d'accord à Karine quand elle dit pour le coup, on est à 4 millions, on se dit, ne peut-on pas donner de l'info ? Non, ça ne marche pas comme ça. La règle n'est pas comme ça, donc on ne peut pas se le permettre.

Juste pour revenir, pourquoi ces huit réserves importantes ? Le point majeur, c'est tout de même que c'est une maladie rare et ce n'est pas facile comme dossier. Si on vous propose un dossier maladie rare, faites-le avec quelqu'un qui en a déjà fait, parce que ce n'est pas facile. Vous allez souffrir un peu plus que pour les dossiers plus traditionnels, même s'il n'y a pas de dossiers faciles. Les dossiers maladies rares sont particuliers.

Le deuxième point, la cheffe de projet en a parlé quand elle a parlé des coûts, mais elle n'avait plus de temps, donc elle a fait vite. Je vais vous le dire de façon très simpliste et c'est assumé. Il y a tout de même un truc particulier dans cette maladie, c'est que tout est basé sur ce fameux score mFARS ou SARA, c'est quasiment équivalent. Ce score, la cheffe de projet vous l'a dit, mais vite, or c'est tout de même important quand on fait de la modélisation, c'est que pour un même niveau de score, vous allez avoir des états de santé qui vont être différents. C'est assez inhabituel, quand on fait de la modélisation économique ou généralement, on raisonne en discret. On dit, il a tel état de santé, tel coût, telle utilité. Là, c'est un peu plus compliqué, ils ont essayé de traduire ça parce qu'ils fonctionnent en continu. Quand je dis eux, c'est les cliniciens, pour eux, ce score, c'est un des scores habituels, mais le faire rentrer dans un modèle médico-économique, ce n'est pas facile.

Vous imaginez bien le nombre d'hypothèses – vous l'avez tout de même vu – qu'il faut faire pour passer de quelque chose de continu qui reflète différents états de santé – parce que c'est une maladie qui est complexe avec des états cliniques très différents – à quelque chose, finalement, qui est plus ou moins discret avec un état de santé, une utilité, un coût. En faisant ça, ils simplifient au maximum leur modèle et donc en simplifiant, il y a une incertitude qui est majeure. On se pose même la question de se dire : le modèle reflète-t-il vraiment la complexité ? Est-il encore en mesure de refléter la complexité de la pathologie ? On n'en est plus trop sûr à la fin, et c'est un peu ça qui pose un problème.

Ce n'est pas de leur faute, encore une fois, c'est juste que la maladie est compliquée, qu'en plus les données sont peu nombreuses. Moins on a de données, moins on peut faire des modèles complexes. Du coup, ces réserves traduisent cette incertitude qu'on a.

Pour répondre un peu et comme je redonne la parole à Sandy, quand elle dit peut-on finalement un peu borner ça ? En fait, on ne peut pas. Enfin on peut, mais ce n'est pas notre rôle. Parce que comme ce sont des paramètres qui sont différents, il faudrait qu'on fasse un scénario en disant si tous les paramètres sont très négatifs, qu'est-ce que ça donne ? Et si au contraire on prend plutôt quelque chose de très positif, qu'est-ce que ça donne, mais ça réimplique de demander au chef de projet de remanipuler le modèle et ce n'est pas toujours très facile de le faire. Ça induit un risque et c'est une responsabilité collective qu'on prend de ressortir des chiffres à partir d'un modèle qu'on n'a pas construit, même si on est censé tout de même un peu connaître comment il fonctionne. C'est quelque chose qu'on n'aime pas bien faire globalement. Autant sur les analyses d'impact budgétaire, on peut tester, parfois juste changer un coût, mais quand là, c'est sur des tas de paramètres, c'est un risque pour nous de se mouiller alors que peut-être qu'on va sortir des résultats qui sont faux.

M^{me} TUBEUF.- Ils auraient pu le faire eux-mêmes.

M^{me} MARGIER.- Ils auraient pu le faire, mais on leur a posé des questions.

M^{me} TUBEUF.- Ce n'est pas à nous de le faire.

M^{me} CHEVREUL.- Hassan, puis Antoine.

M. SERRIER.- Je rebondis. C'est un point extrêmement important parce que vous verrez, quand vous aurez un dossier, la particularité par rapport à l'habitude qu'on a de lire une évaluation médico-économique, c'est que là, ils doivent nous expliquer dans un détail extrême, le dossier est extrêmement détaillé.

Du coup, on va avoir une analyse de sensibilité très détaillée, sauf qu'en fait, malgré tout, cette analyse de sensibilité, tous les paramètres, analyses de scénario, tournent autour des valeurs de

référence. Quand on a un paramètre, deux paramètres avec lesquels on n'est pas d'accord, il faudrait ne serait-ce qu'un seul qui change. Il faudrait que l'analyse de référence soit celle-ci et que toutes les analyses de sensibilité changent.

C'est pour cela que l'étape de l'échange technique est extrêmement importante, c'est presque le plus important, c'est d'avoir en analyse de référence les paramètres que l'on aimerait. Là, j'ai cru comprendre qu'il y a des points qui avaient été demandés, qui n'ont pas été modifiés en échange technique et c'est dommage. Parce que peut-être que s'ils avaient mis en analyse de référence les choix conservateurs par exemple ou autres, peut-être qu'ils n'auraient pas eu tout ça.

M^{me} CHEVREUL.- Après, la vraie question qui se pose, c'est de savoir pourquoi, quand on leur a conseillé à l'échange technique de faire d'une façon, ils ne changent pas, justement. Pour le coup, cela nous rend toujours un peu suspicieux, mais ce n'est peut-être pas normal, mais c'est comme ça. Antoine.

M. BENARD.- Je voudrais revenir sur l'incertitude du modèle. Il y a deux éléments principaux : la différence de coûts et la différence d'efficacité, je reviendrai sur la différence d'efficacité après. Sur la différence de coûts, pour moi, il y a assez peu d'incertitude. Il y a le coût du médicament qui est donné ou pas donné. Quand il est donné, cela coûte [REDACTED] euros par mois, donc [REDACTED] euros par an. Au bout de 20 ans, cette différence de coût s'accumule.

Après, il peut y avoir une incertitude sur le coût des événements indésirables, le coût des états de santé, qui peut être un peu hétérogène, mais le coût du médicament est tellement élevé que sur les coûts, pour moi, il n'y a pas tellement d'incertitude.

Après la question de la différence d'efficacité se pose. Effectivement, il y a de l'incertitude, il y a des hypothèses fortes qui sont faites. Globalement, on a vu sur l'essai clinique la différence, l'efficacité du traitement n'est pas très élevée. C'est peut-être cohérent avec les six mois de plus en bonne santé dans le groupe traité à un horizon temporel de 20 ans, mais on est dans une différence d'efficacité qui n'est pas très élevée. Le coût est tellement élevé et peu incertain qu'à mon avis, on peut tout de même pencher sur un avis plus tranché que de dire simplement il y a une incertitude globale et on ne se prononce pas. J'ai l'impression qu'on va dans cette direction.

M^{me} CHEVREUL.- Je suis assez d'accord, mais avant qu'on parle, j'aimerais bien avoir l'avis des cheffes de projet sur cette approche.

Une cheffe de projet, pour la HAS.- C'est intéressant, on sait que sur les coûts, on a le coût du produit qui est extrêmement important et, cela limite notre incertitude. Après, d'un point de vue méthodologique, déjà on a une première approche. La réserve sur les coûts que vous avez là vient aussi essentiellement de la cohérence. Le problème, c'est que cette approche catégorielle qui est

proposée aujourd'hui dans notre modèle où on nous explique qu'on a une évolution continue de la pathologie modélisée à travers le score mFARS, vient faire un peu un non-sens avec ce qui était proposé dans la modélisation de la pathologie. Il me semble que dans notre conclusion, on l'avait notamment précisé sur l'intégration de la qualité de vie et des données d'utilité.

Sur l'incertitude globale majeure. Certes, les coûts sont une variable sur laquelle on pourrait dire qu'on a moins d'incertitudes que la qualité de vie ou l'efficacité. Après, dans le modèle, au-delà de cela, il y a tout de même d'autres dimensions qui rentrent en considération. On a tout de même la transposabilité de la population simulée. C'est un premier point sur lequel aujourd'hui, ce qui a été proposé dans le modèle, on ne sait pas si c'est transposable à la population qui sera traitée en pratique clinique. On ne sait même pas si les données qui sont utilisées dans le modèle sont finalement transposables à la population qui est simulée dans ce modèle.

On a également des questions, comme le disait Jennifer, sur la modélisation de la pathologie en tant que telle. On n'est pas sûr d'avoir, dans le modèle utilisé, pu refléter suffisamment la complexité de cette pathologie.

L'incertitude n'est pas tant sur l'idée de dire qu'on croit dans les résultats de coûts ou qu'on ne croit pas, et qu'on croit dans les résultats d'efficacité ou non. Elle vient vraiment dans l'idée que par cette approche complète, on a ces différents points et cette incertitude sur différents aspects qui font qu'aujourd'hui, dans la modélisation telle qu'elle est faite, on ne croit pas dans les résultats. On a une approche beaucoup plus globale.

Dans le projet d'avis, on a une proposition de conclusion dans laquelle on retrace les différents points sur lesquels on a de l'incertitude. On voit bien que c'est le modèle, l'intégration des données cliniques, les hypothèses à long terme, le maintien de l'effet traitement, la durée de traitement. C'est tout cela qui fait une accumulation qui fait vraiment qu'il y a trop d'incertitudes sur tous ces aspects pour qu'on croie réellement dans une partie ou une autre partie des résultats.

M. BENARD.- Nos deux points de vue ne sont pas incompatibles. Je ne remets pas en cause le fait qu'il y ait une incertitude globale sur le modèle, simplement je dis que cette incertitude ne va pas jouer jusqu'au point où l'on peut considérer à un moment donné que le traitement est efficient dans le système de point.

M^{me} CHEVREUL.- C'est aussi la raison pour laquelle vous n'avez pas proposé une incertitude globale sur l'impact budgétaire. Effectivement, sur le coût, on a un peu d'incertitude, mais on a beaucoup plus d'incertitude sur la modélisation et la prévision qu'on peut faire de l'efficacité à moyen terme.

Cela vous va-t-il si on prend un café histoire de le digérer avant de présenter l'avis justement parce que cela va aider peut-être à échanger entre nous de façon informelle et arriver sur l'avis ? Cela vous convient-il si on fait comme cela ?

Si tout le monde est prêt, on va avancer. La procédure d'après c'est que nous nous mettions d'accord sur un avis. Je vais laisser la parole aux cheffes de projet pour faire la proposition d'avis qui reprend ce qu'elles ont dit et éventuellement ce qu'on a dit pour qu'on se mette d'accord sur ce qu'on pourra proposer.

Une cheffe de projet, pour la HAS.- En termes de propositions de conclusions, encore une fois c'est une proposition. Pour l'efficience de ne pas présenter les résultats pour raison d'une incertitude globale majeure, avec dans la proposition de conclusion de détailler les différentes incertitudes, c'est ce que vous retrouvez sur la diapositive, sur la transposabilité de la population simulée, la modélisation de l'évolution des patients, l'efficacité relative, la qualité de vie et l'effet traitement. Dans chaque paragraphe, il est détaillé les différentes incertitudes.

Ensuite, dans l'impact budgétaire, en revanche, de présenter les résultats en rappelant que certaines hypothèses n'impactent pas de manière significative les résultats de l'impact budgétaire, qui permettent de les présenter.

Si vous n'êtes pas forcément familier avec le format du projet d'avis, il y a d'abord une section sur la conformité méthodologique où l'on rappelle les réserves, ensuite une première pré-conclusion en ce qui concerne l'efficience, on a le même schéma pour l'impact budgétaire, ensuite un paragraphe sur la conclusion générale où l'on essaie de donner des messages de manière moins technique et de manière plus générale pour le lecteur. Ensuite, une section de données complémentaires est aussi proposée.

M^{me} CHEVREUL.- Vous voulez peut-être le balayer tout de même. Normalement, il vous a été envoyé, pour que vous ayez le temps de le lire avant d'avoir eu cet échange. Il faut préciser ce que tu disais Hassan, c'est que si effectivement on va sur une incertitude globale, alors on ne représente plus les résultats et on ne représente peut-être plus l'ordre du 4 millions d'euros par QALY, c'est ça ?

Ce qui est vraiment important dans ce qu'on vient de dire, c'est ce que tu disais à ce propos et ce que disait Antoine, en disant : bon d'accord il y a une incertitude, mais elle ne doit pas être si importante que ça, puisque de toute façon, sur l'évaluation du coût, ■■■ % du coût c'est le coût du médicament, et les incertitudes qu'on a c'est sur l'efficacité. C'est ce que tu disais. Là, il nous reste à savoir si on veut toujours aller sur une incertitude globale majeure qui fait qu'on ne présente pas, ou si on dit qu'il y a une incertitude qui est très forte avec telle et telle réserve, et on présente tout de même.

Je vois qu'Hassan lave la main, donc ce sera le premier à répondre à la question.

M. SERRIER.- Non, je ne vais pas répondre à la question, mais juste une précision qui a son importance dans ce dossier. C'est peut-être moins important que dans d'autres, puisqu'il y a une ASMR V qui a été obtenue. A priori, cela ne devrait pas impacter derrière la fixation.

M^{me} CHEVREUL.- La proposition qui est faite suite à nos échanges, c'est d'aller vers l'incertitude globale majeure à cause du nombre de réserves importantes qui a été mis sur l'évaluation de l'efficience, que ce soit justement sur l'évaluation de l'efficacité, de l'utilité, du coût, et autres. Y a-t-il une contre-proposition par rapport à cela, par rapport à ce qui s'est dit ?

M. SPATH.- Justement pour conforter, il faut dire aussi aux nouveaux que les industriels qui rédigent les dossiers regardent bien les dossiers que nous avons évalués. Une fois qu'on a pris telle décision, pour un autre dossier qui est aussi difficile pour ne pas dire mauvais, mais difficile et complexe, comme Jennifer et tout le monde l'a dit, on reprocherait : mais pourquoi pour celui-là vous n'avez pas mis d'incertitude globale majeure et pour vous, vous la mettiez. Il faut faire très attention à être cohérent entre ce qu'on dit sur les dossiers.

M^{me} CHEVREUL.- En général, c'est Hassan qui dit cela. Je vois qu'en fait, maintenant, vous êtes d'accord. Je sais que c'est toi, Hassan, qui a été le premier à discuter éventuellement qu'on aborde l'incertitude globale majeure. Quelle est ta position maintenant ?

M. SERRIER.- Encore une fois, je suis très pragmatique et cela me gêne quand on a une information importante qu'on peut fournir et qu'on ne la fournit pas. Je reste persuadé, Antoine l'a mieux dit que moi, mais que le ratio est élevé, même s'il y a une incertitude globale majeure, que le produit n'est pas efficient. Maintenant, encore une fois, c'est une ASMR V. Cela veut dire que cela ne va pas rentrer en ligne de compte. De manière pragmatique, cela ne changera rien. Je me battraï moins là-dessus sur ce dossier.

M^{me} CHEVREUL.- D'accord. Pour le coup, on peut rester dessus. J'ai juste une information complémentaire : ils ont le dossier. On ne leur présente pas dans l'avis, mais ils ont bien l'évaluation qui a été faite par l'industriel.

M. BEAUFILS, pour la HAS.- Non, ils ont toutes les informations dans l'avis. C'est simplement qu'ils savent que nous avons invalidé les choses et qu'ils ne vont pas les utiliser.

M^{me} CHEVREUL.- Mais dans le coup, sont-ils informés tout de même de ce qu'avait proposé l'industriel ?

M. BEAUFILS, pour la HAS.- Oui. Il y a une partie au début sur le contexte avec les revendications initiales de l'industriel, puis l'analyse critique. Ils savent que l'industriel a revendiqué ci, a revendiqué un RDCR de temps, etc. Après, il y a marqué à côté que c'est non retenu par la CEESP.

M^{me} CHEVREUL.- Ils ont tout de même cette information.

Une cheffe de projet, pour la HAS.- Les résultats sont présentés dans une section dans l'avis un peu plus tard dans un tableau vous retrouvez tous les résultats. C'est juste qu'on ne les reprend pas dans la conclusion. Sinon, ils sont bien mentionnés dans l'avis.

M. BENARD.- Je ne suis peut-être pas le seul à me poser la question, mais je ne connais pas bien l'articulation au sein des commissions. S'il y a ASMR V à la Commission de la transparence, qu'est-ce que cela a comme conséquences dans l'évaluation du produit ? Cela veut dire qu'il n'y a pas de conséquences de l'avis de la CEESP ?

M^{me} CHEVREUL.- Non, je pense que c'est plus comment le CEPS va négocier le prix. Ça n'a pas de conséquences sur notre avis. Globalement, ils ont plusieurs informations et ils s'en servent pour la tarification.

Un chef de projet, pour la HAS.- Dans le process de négociation des prix, le CEPS utilise nos avis économiques. Il utilise les prix des cinq pays du corridor européen. Il utilise les achats hospitaliers. Ils utilisent pas mal d'informations. L'évaluation économique est une variable comme une co-variable dans une régression, une variable aujourd'hui d'ajustement. L'ASMR figure déjà comme un facteur défini de manière légale depuis l'adoption de l'évaluation économique en France des produits de santé en 2013.

M^{me} CHEVREUL.- L'étape d'après, à moins que vous ne vouliez rediscuter et faire des contre-propositions à ces avis, c'est de voter l'avis.

Une cheffe de projet, pour la HAS.- L'impact budgétaire, êtes-vous d'accord pour présenter les résultats de l'impact budgétaire ?

M^{me} CHEVREUL.- Je suis d'accord. Quand je disais de l'avis, c'était l'avis complet, aussi celui de l'impact budgétaire. Cela convient-il à tout le monde ?

Une intervenante.- Si on est d'accord sur l'avis, cela ne veut pas dire qu'on est d'accord sur le prix demandé pour la boîte.

M^{me} CHEVREUL.- Non.

L'intervenante précédente.- D'accord, c'était pour savoir. Comme base de négociation soit, mais je crois que quand il y a le pire des ASMR, le prix doit être très bas. Enfin, il y a une espèce de règle tacite.

M^{me} CHEVREUL.- Cela reste à la main du CEPS.

Une cheffe de projet, pour la HAS.- Ce qu'on fournit dans l'avis, ce sont les analyses de sensibilité avec des baisses de prix. Ils ont les résultats de l'avis avec différentes simulations sur des prix alternatifs.

M^{me} CHEVREUL.- Si tu veux peut-être remettre la synthèse. Gaëtan.

M. DUPORT.- Peut-être que je ne le vois pas dans la conclusion. Juste, j'ai l'accord-cadre CEESP-LEEM sous les yeux, pour éviter tout le litige. Il y a écrit : « Trois situations principales peuvent conduire à ne pas pouvoir établir ou questionner les conditions d'efficience d'un produit ». On est en train de discuter de ça. Il y a écrit « Lorsque l'incertitude globale générée par un ensemble de choix méthodologiques ou par les données entrées dans le modèle s'avère très forte et qualifiée comme telle dans l'avis CEESP ». Est-ce le cas ici ?

M^{me} CHEVREUL.- C'est pour cela qu'on va vous remettre à l'écran.

M. DUPORT.- Je ne l'ai pas vu.

M. BEAUFILS, pour la HAS.- Non, mais il n'y a aucun souci. Nous, on utilise le terme « incertitude globale majeure ». Eux, ils utilisent le terme « incertitude très forte ». C'est très clair dans la tête des entreprises et du CEPS que ça veut dire la même chose. Nous utilisons bien « incertitude globale majeure » qui a cette traduction dans l'accord-cadre. C'est un peu étonnant, mais c'est comme ça.

M^{me} CHEVREUL.- C'est le mapping. Nous vous soumettons cette conclusion. Si personne ne souhaite la changer, il faudra donc que nous votions. Êtes-vous prêts à voter ou y a-t-il quelqu'un qui a des remarques ou des propositions à faire avant ? Samantha, je vous laisse la main.

(Il est procédé au vote par appel nominatif.)

M^{me} FERNANDES, pour la HAS.- J'ai 18 pour et 5 abstentions.

M^{me} CHEVREUL.- Très bien. Merci. Merci aux cheffes de projet, pour ce dossier qui est tout de même un des plus complexes qu'on ait vus dernièrement, en tout cas pour mon expérience qui n'a commencé qu'au mois de mars.