

Commission Évaluation Économique et Santé Publique

13 avril 2021

TRANSCRIPTION DES DÉBATS

AVERTISSEMENT

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

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

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

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

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

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

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

1. Avis économique ECO-EFFI 433 Calquence (AstraZeneca)

M^{me} PARIS. – On peut passer à l'examen de l'avis économique de Calquence.

Un chef de projet, pour la HAS.- Merci, bonjour à tous. Je vais partager mon écran. On voit aujourd'hui Calquence, l'acalabrutinib, dans le traitement de la leucémie lymphoïde chronique.

C'est un traitement d'une hémopathie maligne caractérisée par l'accumulation et la prolifération de lymphocytes B d'aspect mature avec envahissement sanguin et médullaire. On estime qu'il y a un peu plus de 4 500 cas incidents en France en 2018, et cela représente 1 % des cas incidents de cancer en France. La prise en charge et la surveillance pour les stades précoces et asymptomatiques sont une mise en œuvre d'un traitement lorsque la maladie atteint un stade avancé, le stade B et C et/ou si la maladie est symptomatique ou active.

Le produit est un inhibiteur sélectif de la tyrosine kinase de Bruton de seconde génération. C'est un traitement par voie orale, ce qui le distingue d'autres traitements disponibles dans l'indication, qui se font par voie injectable. Le produit a eu une indication en novembre 2020, soit en monothérapie ou en association avec l'obinutuzumab dans le traitement des patients adultes atteints d'une leucémie lymphoïde chronique non précédemment traitée, donc en première ligne, ou en deuxième ligne, en monothérapie.

L'industriel revendique une primo inscription en ville et à l'hôpital, avec un SMR important, une ASMR III en première ligne en monothérapie et à l'inverse, une ASMR IV en première ligne en association et en deuxième ligne en monothérapie.

Le prix revendiqué et de [REDACTED] euros hors taxes pour 60 gélules et de RDCR revendiqué est de 331 158 euros par QALY *versus* l'ibrutinib. Cette revendication de RDCR concerne la première ligne en monothérapie.

L'impact attendu est de presque 1 100 patients incidents en 2020 dans l'indication de l'évaluation que vous allez le voir après. C'est une indication restreinte par rapport au remboursement et à l'AMM. Le chiffre d'affaires revendiqué est de [REDACTED] d'euros par an, toutes indications confondues.

Vous avez ici ce que représente l'évaluation économique par rapport au périmètre du remboursement. L'évaluation économique est ciblée sur l'ASMR III revendiquée. La première ligne est en monothérapie chez des patients qui sont soit inéligibles à la fludarabine à haute dose,

soit chez des patients qui ont un statut génétique de mauvais pronostic, c'est-à-dire une délétion 17p ou une mutation TP 53.

Le remboursement est revendiqué à la fois en monothérapie en première ligne ou en bithérapie en première ligne, ou en deuxième ligne en monothérapie. L'AMM est encore plus large que le périmètre de remboursement demandé, puisqu'il n'y a pas de restriction chez les patients en fonction du statut génétique ou des difficultés à la fludarabine.

Les données cliniques disponibles sont uniquement celles issues de l'essai ELEVATE TN qui a inclus des patients qui étaient inéligibles à fludarabine à haute dose, et qui a ensuite stratifié des patients en fonction du statut mutationnel. Vous voyez que 535 patients étaient inclus, parmi lesquels 73 patients, un peu moins de 14 % avaient un statut mutationnel spécifique. La randomisation s'est faite selon trois groupes :

- le comparateur obinutuzumab + chlorambucil
- l'obinutuzumab + acalabrutinib, l'association de l'acalabrutinib en monothérapie qui fait l'objet d'évaluation économique.

Le cross-over était possible après progression pour les patients qui étaient sous chimiothérapie, et pouvaient basculer principalement sur l'acalabrutinib en monothérapie. Le traitement était soit jusqu'à la progression, soit avec un nombre de cycles prédéfinis pour la chimiothérapie.

Les critères de jugement de l'essai clinique étaient hiérarchisés :

- En premier critère l'évaluation de la bithérapie, pour laquelle une ASMR IV est revendiquée *versus* le Groupe chimiothérapie. Vous avez ici le résultat à 24 mois, avec un HR de 0,10. Le critère était satisfait.
- Le second critère de jugement hiérarchisé, c'est la monothérapie *versus* la chimiothérapie. Là encore, on observe un résultat significatif avec un HR de 0.20.
- Ensuite, nous avons deux critères de taux de réponse objective, d'abord pour la bithérapie, ensuite pour la monothérapie. Autant le premier résultat est significatif, autant le deuxième résultat n'est pas significatif, ce qui arrête ici la hiérarchie des critères de jugement.

Vous pouvez observer que les critères de survie globale, qui étaient des critères secondaires plus loin dans la hiérarchie, n'apparaissent pas comme significatifs : le critère hiérarchisé n'est pas atteint, les médianes de survie ne sont pas atteintes non plus. Nous n'avons donc pas de données matures sur la survie globale dans cet essai clinique.

Concernant l'analyse de l'efficience, les choix structurants sont assez classiques sur le type d'analyse. On a une analyse coût utilité, coût efficacité. La perspective est collective. L'industriel a déposé des analyses complètes à cinq ans, dix ans et quinze ans et nous avons retenu en analyse de référence l'horizon temporel de dix ans, qui correspond à l'arbitrage entre le niveau d'incertitude et la pertinence de l'évaluation.

Deux questions se posent concernant la population d'analyses et les comparateurs. Pour la population d'analyses en première ligne de traitement, on a d'une part les patients inéligibles à la fludarabine à haute dose, et d'autre part, les patients qui ont une délétion 17p ou une mutation TP 53. L'industriel fait une analyse de référence dans l'ensemble de cette population et une analyse exploratoire chez les patients qui ont un statut génétique spécifique.

Les comparateurs inclus sont

- ibrutinib,
- obinutuzumab en association au chlorambucil,
- bendamustine + rituximab,
- chlorambucil + rituximab
- acalabrutinib + obinutuzumab, qui est la bithérapie de l'acalabrutinib.

Sont exclus :

- idelalisib + rituximab
- venetoclax + obinutuzumab

Je vais aller plus en détail sur la question des populations. Vous avez ici les comparateurs qui sont recommandés soit par la commission de la transparence, soit par le groupe d'experts professionnels, le groupe FILO, en fonction du profil des patients. D'une part, les inéligibles à la fludarabine, et d'autre part les patients qui ont eu une délétion 17p ou une mutation TP 53, et la pertinence d'inclure ces produits comme comparateurs selon les experts interrogés par l'industriel.

Vous pouvez observer ici que les comparateurs sont différents selon des sous-groupes, ce qui nous permet de conclure, dans l'analyse critique, à la nécessité d'avoir des analyses en sous-population puisque les comparateurs sont différents. Les pronostics sont différents également puisque la présence d'une anomalie génétique entraîne un plus mauvais pronostic. Il s'agissait en outre d'un critère de stratification dans l'essai clinique.

Nous proposons une réserve importante en l'absence d'analyse chez les patients avec des délétions 17p ou une mutation TP 53 sans autre motif d'inéligibilité à fludarabine, puisque l'ensemble des patients inclus dans l'essai clinique étaient en premier lieu inéligibles à la fludarabine, ensuite, en second lieu, éventuellement avec une délétion 17p ou une mutation TP 53. Il n'y avait pas de patients qui auraient eu la délétion ou la mutation sans être par ailleurs inéligibles à la fludarabine.

Concernant les comparateurs, vous avez ici à nouveau la liste des comparateurs, et la question se pose autour du venetoclax + obinutuzumab qui a été exclu des comparateurs de l'analyse de l'efficacité, sur l'argument que ce produit n'est pas recommandé par la commission de la transparence. L'industriel a eu une autorisation de mise sur le marché dans l'indication, mais il n'a pas à solliciter le remboursement auprès de la Commission de la transparence. Il est néanmoins recommandé par le groupe d'experts FILO dans ses recommandations de 2020. Il est recommandé aussi, selon les experts sollicités par l'industriel. Nous vous proposons une réserve importante sur l'exclusion de ce comparateur dans l'analyse.

Concernant la modélisation, la comparaison entre les traitements repose sur une méta-analyse en réseau qui a été menée chez les patients inéligibles à la fludarabine sans prise en compte spécifique du statut mutationnel. Vous avez ici la totalité du réseau qui relie les traitements : l'acalabrutinib en haut à droite, qui est le traitement évalué. Ici, vous avez tous les comparateurs qui apparaissent encadrés en bleu. Les comparateurs en trait plein sont des comparateurs inclus dans l'analyse et que vous avez le venetoclax + obinutuzumab qui n'est donc pas inclus.

Le comparateur le plus utilisé, c'est la bendamustine + rituximab, qui représente 38 % des patients traités. Et vous avez l'ibrutinib qui est le comparateur qui ressort sur la frontière d'efficacité, mais qui est relativement peu utilisé chez les patients inéligibles à la fludarabine à haute dose parce qu'ils ne représentent que 2 % des traitements.

La comparaison indirecte sous la forme d'une méta-analyse en réseau s'appuie sur l'analyse de la vérification de l'hypothèse des risques proportionnels. C'est le fait de vérifier que l'effet est constant dans le temps au sein des essais cliniques, ce qui permet ensuite d'utiliser le hazard-ratio estimé dans la méta-analyse pour comparer les traitements dans la modélisation. On observe que dès l'essai clinique princeps, l'essai clinique pivot qui est utilisé pour commencer à relier le produit au réseau, l'hypothèse des risques proportionnels n'est pas vérifiée. Nous savons que l'effet relatif des traitements entre eux n'est pas constant dans le temps, ce qui empêche l'utilisation des HR issus de la méta-analyse pour l'estimation de l'effet relatif des traitements.

En outre, la méta-analyse présente d'autres limites. D'une part, pour certains essais cliniques, l'évaluation est bien disponible selon le comité indépendant, qui est le critère de jugement qui est privilégié. Mais pour certains essais, il n'est disponible que selon l'évaluation faite par l'investigateur dans un essai en ouvert. On observe dans l'essai acalabrutinib pour lequel les deux critères de jugement selon qu'ils sont observés par le comité indépendant ou par l'investigateur, sont disponibles et entraînent une différence assez sensible du hazard-ratio, ensuite du RDCR, si jamais on simule les résultats à partir de l'un ou l'autre des critères.

Le cross-over était possible dans certains essais, mais pas dans d'autres, ce qui ne pose pas de problème pour la survie sans progression, puisque les patients ne pouvaient changer de traitement qu'après progression. Par contre, cela pose un souci pour l'estimation de la survie globale puisque dans certains cas, on n'évalue que le premier produit, et dans d'autres cas, l'ensemble des produits successifs.

Enfin les caractéristiques des patients pouvaient varier d'un essai à l'autre dans la méta-analyse, dans des proportions relativement importantes, notamment concernant le statut mutationnel, qui peut avoir un impact sur la réponse au traitement.

L'industriel a néanmoins retenu le hazard-ratio issu de la méta-analyse standard, en argumentant que la majorité des études respectaient l'hypothèse des risques proportionnels. Il me semble que c'est tout de même problématique dans la mesure où elle n'est pas respectée, en particulier dans l'essai de départs du réseau.

Il a été exploré différentes approches alternatives, d'une part, la méta-analyse à polynômes fractionnaires qui n'est pas conclusive. Les modèles ne convergent pas, et on ne peut donc rien en déduire. D'autre part, l'approche par *restricted mean survival time*, qui est très incomplète puisque l'industriel ne fait que comparer les hazard-ratios estimés selon cette méthode dans les essais princeps, aux hazard-ratios estimés directement dans les essais selon l'investigateur. Nous n'avons aucune analyse qui irait jusqu'au bout, c'est-à-dire en estimant que les hazard-ratios dans les méta-analyses et pas seulement dans les essais au départ, ensuite en simulant les résultats dans la modélisation. Cela aurait permis éventuellement de mener cette approche alternative jusqu'au bout.

L'industriel discute également une approche selon une MAIC qui aboutirait à des comparaisons deux à deux, ou une comparaison limitée à l'essai clinique, ce qui représenterait des limites très fortes dans l'analyse de l'efficacité.

Enfin l'industriel applique le hazard-ratio pendant 28 mois, c'est-à-dire la durée médiane des essais cliniques de la méta-analyse, puis applique un hazard-ratio de 1. Cela revient à faire l'hypothèse qu'au-delà des 28 mois de l'essai clinique, l'effet traitement est similaire entre les

bras de traitement, alors que l'on sait, puisque l'hypothèse des risques proportionnels n'est pas vérifiée, que cet effet traitement n'est pas du tout le même selon le produit dans le temps.

L'industriel explore également la possibilité de faire une méta-analyse en réseau pour les patients porteurs de mutation ou délétion. Etant donné le faible nombre de patients concernés, il regroupe les patients qui reçoivent de l'acalabrutinib, en monothérapie ou en bithérapie, et il cherche à le comparer à l'ibrutinib, qui est le comparateur de référence dans cette population, puisqu'il représente 74 % des patients traités dans cette sous-indication, dans cette sous-population.

L'analyse exploratoire n'est pas conclusive, parce qu'il n'y a que 73 patients qui sont inclus dans les deux bras de traitement de l'essai clinique. On n'observe que 7 évènements de progression, ce qui n'est pas du tout suffisant pour mener une comparaison indirecte. Nous vous proposons donc une réserve importante du fait de l'absence d'évaluation chez les patients, qui sont à la fois inéligibles à la fludarabine et porteurs de délétion ou de mutation.

Concernant l'extrapolation temporelle, l'industriel a testé plusieurs fonctions d'extrapolation qui sont illustrées sur le graphique. Vous voyez que cette exploration conduit à une très grande variabilité dans les résultats à l'horizon temporel des 300 mois, font de 0 à 50 % de survie sans progression. L'industriel retient, en analyse de référence, la fonction gamma généralisée, qui est la fonction grise sur le graphique, et en analyses de sensibilité, la fonction exponentielle qui minimise les critères statistiques, c'est la courbe en rouge, et la fonction Gompertz qui est la moins favorable à l'acalabrutinib, qui est la fonction jaune sur le graphique. Cette approche n'appelle pas de réserves en particulier, mais illustre la forte incertitude qui entoure les styles d'extrapolation des résultats.

Concernant l'estimation de la survie globale, pré-progression, les données disponibles sont

- d'une part les données des essais cliniques, mais ce n'est généralement pas un critère de jugement principal dans les essais,
- les données in mature de l'essai clinique, de l'essai pivot, avec des différences qui n'apparaissent pas du tout significatives dans la méta-analyse en réseau, au-delà de toutes les limites que j'ai déjà évoquées sur la méta-analyse.

L'approche de l'industriel, qui avait été évoqué en rencontre pré-dépôt et qui a été redemandé en échange technique, c'était d'appliquer un HR de 1, c'est-à-dire de faire l'hypothèse qu'il n'y a pas d'effet relatif du traitement sur la mortalité en l'absence de preuves statistiquement significatives. Mais l'industriel n'a pas retenu cette approche. Il applique le hazard-ratio de la survie sans progression à la survie globale pendant les 28 premiers mois de la modélisation.

Pour justifier son choix, il s'appuie sur deux essais cliniques qui sont faits chez des comparateurs : l'obinutuzumab et l'ibrutinib, dans lequel on observe à la fois un effet significatif sur la survie sans progression et sur la survie globale. Il utilise cet argument en justifiant le fait que l'on observe bien, quand il y a un effet significatif sur l'un, qu'il y a aussi un effet significatif sur l'autre.

Vous observerez néanmoins que cet effet significatif ne va pas toujours dans le même sens, puisque le hazard-ratio est plus élevé pour la survie globale que sur la survie sans progression dans l'essai sur obinutuzumab. C'est l'inverse sur l'ibrutinib ce qui ne permet pas d'estimer finalement une corrélation entre la survie sans progression et la survie globale.

L'industriel dans une analyse en scénario utilise les hazard-ratios issus de la méta-analyse en réseau. Vous avez dans le tableau qui est à l'écran une comparaison des hazard-ratios de la survie globale par rapport à la survie sans progression. Vous voyez que les hazard-ratios ne sont pas du tout dans le même ordre de grandeur. Ils vont parfois dans un sens, parfois dans l'autre, ils sont parfois plus élevés, parfois plus faibles.

Je passerai rapidement sur ce point, puisque j'ai déjà par ailleurs présenté toutes les limites de la méta-analyse en réseau. Nous proposons une réserve majeure sur l'estimation de la survie globale avant progression.

Concernant le traitement de seconde ligne. La modélisation est un modèle de Markov, donc on démarre par l'application de courbes de Kaplan-Meier, sur la survie sans progression. Lorsque le patient progresse, il passe en deuxième ligne de traitement, pour lesquels on réapplique les courbes de Kaplan-Meier post-progression pour estimer la survie globale, cette fois-ci.

Dans le tableau qui est présenté, vous avez :

- le traitement reçu par le patient en première ligne,
- ce qui est envisageable selon les recommandations en deuxième ligne, dans la deuxième colonne, avec la justification clinique qui explique pourquoi le patient recevrait l'une ou l'autre des deuxièmes lignes,
- puis la dernière colonne, les données disponibles finalement et l'extrapolation effectivement réalisée par l'industriel dans la modélisation pour simuler la deuxième ligne.

Vous pouvez voir que, de façon très synthétique, si le patient a reçu en première ligne un inhibiteur de BTK, l'effet du traitement de deuxième ligne sera simulé à partir de l'essai MURANO, essai qui évaluait rituximab + venetoclax *versus* bendamustine + rituximab, avec la limite que la population de cet essai clinique n'avait pas reçue d'inhibiteurs de BTK en première ligne. Nous n'avons pas une population superposable à celle de la modélisation. Et surtout les résultats de

l'essai MURANO n'étaient pas du tout significatifs sur la survie globale. Cela ne permet pas d'estimer un effet traitement sur la survie globale dans la modélisation.

Lorsque le patient, à l'inverse, recevait une chimio immunothérapie en première ligne, en deuxième ligne, l'industriel prend les résultats de son propre essai clinique après cross-over pour les patients qui ont reçu la chimio-immunothérapie en première ligne et qui switcherait sur acalabrutinib ou ibrutinib en deuxième ligne. Cela reprend les données de son essai clinique dont on a vu que les résultats n'étaient pas du tout significatifs sur la survie globale. Nous proposons une réserve majeure sur l'estimation de la survie globale en deuxième ligne de traitement.

Concernant les utilités, les données disponibles sont d'une part les données issues de l'essai clinique, puisqu'il y a eu un recueil du questionnaire EQ5D dans l'essai pivot. L'industriel observe qu'il y a énormément de données manquantes, jusqu'à 60 %, dans l'état post-progression. Cela a abouti à un résultat paradoxal puisque l'on a une utilité qui est plus élevée après progression qu'avant en progression. Cela s'explique probablement par le fait que seuls les patients qui étaient en relativement bonne santé après la progression ont répondu au questionnaire. Cela rend très difficile d'utilisation de ces données.

L'industriel à mener une revue de littérature qui s'avère non concluante, c'est une conclusion que l'on avait observée également dans d'autres dossiers, dans une indication similaire. Il propose de reprendre les valeurs d'obinutuzumab. Ce sont des valeurs qui sont issues d'un avis économique précédent, qui traitait aussi de l'indication de leucémies lymphoïdes chroniques en première ligne.

Pour mémoire, dans ce dossier, une réserve majeure avait été mise sur les utilités en 2015, et a été levée en 2018, suite à la réalisation d'une étude ad hoc de révélation des préférences de la population générale française. Il y avait des scores moyens d'utilité qui étaient associés aux différents stades de la maladie qui avaient été décrits dans la population de l'étude par des vignettes et qui prenaient en compte le fait que le traitement soit par voie orale ou par voie injectable.

L'industriel applique en outre, des décréments d'utilité qui sont liés aux événements indésirables. L'estimation reposait sur une approche de *composites time tradeoff*, en France, sur un échantillon de 100 personnes.

L'industriel reprend ses valeurs déjà publiées dans obinutuzumab pour l'analyse de référence, en prenant la moyenne de traitement oral et injectable. Il n'y a pas d'approche par le biais d'une EQ5D. Il l'utilise les données d'ELEVATE en analyse de sensibilité. Mais cette approche est relativement peu pertinente, puisque l'on a vu qu'il y avait, c'était des scores un peu paradoxaux, étant donné les nombreuses données manquantes.

Nous vous proposons une réserve importante en l'absence de valorisation à partir d'un questionnaireEQ5D. C'est une approche cohérente avec celle qui a été retenue pour obinutuzumab.

En synthèse, sur les principales réserves, nous avons les trois réserves majeures qui portent sur l'effet relatif des traitements, la survie pré-progression et de la survie post-progression, des réserves importantes que j'ai déjà mentionnées et que je ne reprends pas, et d'autres réserves importantes que je n'ai pas détaillées et que je vous présente rapidement.

- Le fait que l'estimation des coûts est très largement fondée sur un avis d'expert, sans aucune mise en regard de données externes
- L'absence d'analyse de sensibilité sur le prix de comparateur.
- L'absence d'équation du RDCR en fonction du prix revendiqué, qui n'a pas été fournie par l'industriel.

Concernant les résultats, je présente rapidement la frontière d'efficience. On a comme un point de référence :

- le chlorambucil et l'obinutuzumab qui sont le comparateur de l'essai clinique,
- l'ibrutinib avec un RDRC d'un peu moins de 200 000 euros par QALY,
- l'acalabrutinib en monothérapie, à 331 158 euros,
- l'acalabrutinib association à un peu plus de 450 000 par QALY.

Vous avez ici la courbe d'acceptabilité multi options avec d'une part l'acalabrutinib en monothérapie qui n'est jamais la stratégie maximisant la probabilité d'être coût efficace, puisqu'elle n'apparaît pas du tout sur les courbes avant plus de 300 000 euros par QALY. Elle est ensuite très vite dépassée par l'acalabrutinib en association.

Concernant l'analyse d'impact budgétaire, le périmètre est celui de l'acalabrutinib en première ligne de traitement, que ce soit en monothérapie ou en bithérapie. L'horizon temporel est de trois ans avec une perspective de l'assurance-maladie obligatoire, et le modèle de l'AIB repose sur une cohorte incidente annuelle. On compare deux scénarios, avec ou sans l'acalabrutinib, encore une fois, toujours en monothérapie ou en association.

L'estimation des coûts repose, comme sur l'analyse d'efficience, sur des dires d'expert. Par contre la valorisation se fait selon les tarifs et non selon les coûts de production. Les analyses de sensibilité portent sur les parts de marché, la taille de la population et le prix de l'acalabrutinib.

Vous avez ici des résultats de l'analyse d'impact budgétaire. On a, à l'horizon de trois ans, un impact cumulé de [REDACTED] d'euros, avec évidemment un impact très fort des coûts d'acquisition des traitements qui est en partie, mais légèrement compensé, par un moindre coût d'administration, puisque l'on passe sur un traitement oral plutôt qu'injectable.

Les réserves sont très similaires avec celles déjà énoncées sur l'efficacité. Nous avons une réserve spécifique, qui est liée au périmètre limité de l'analyse d'impact budgétaire par rapport au périmètre de remboursement, puisque nous n'avons pas du tout d'analyse de l'acalabrutinib en deuxième ligne, alors qu'ils revendiquent le remboursement. On retrouve les réserves sur l'absence d'analyse de sensibilité sur le prix des comparateurs et l'exclusion du venetoclax + obinutuzumab.

En conclusion, je vous présente les principaux éléments. D'une part, le périmètre est restreint par rapport aux remboursements et par rapport à l'autorisation de mise sur le marché. Concernant l'efficacité, le rappel des trois réserves majeures et le fait que malgré ces réserves, l'évaluation de l'efficacité met en avant un certain nombre de sources d'incertitude :

- d'une part sur l'évolution des patients au-delà de l'essai clinique, nous voyons que les fonctions d'extrapolation faisaient apparaître une très forte incertitude sur le devenir des patients à moyen et long terme,
- l'incertitude sur l'effet relatif d'acalabrutinib par rapport aux autres traitements envisageables.
- L'incertitude sur l'effet des traitements de seconde ligne,
- l'incertitude sur les coûts de prise en charge,
- et l'incertitude due à l'absence de validation externe des résultats de la modélisation.

Concernant l'impact budgétaire, il présente quelques réserves importantes. L'impact est estimé à [REDACTED] d'euros à trois ans, soit une augmentation de 15 % du budget de l'assurance-maladie pour la première ligne. Les résultats sont très sensibles au prix d'acalabrutinib et modérément sensibles à la prise de parts de marché.

En conclusion générale –pardon cette conclusion aurait dû apparaître juste avant– l'évaluation réalisée ne permet pas de définir les conditions d'efficacité du produit *versus* les comparateurs avec un degré acceptable d'incertitude.

Enfin, nous demandons des données complémentaires qui portent sur

- l'évolution des patients sous traitement et la corrélation entre survie sans progression et survie globale,

- l'utilité mesurée à partir d'un questionnaire de qualité de vie pour lequel une valorisation de l'utilité est validée,
- enfin, les coûts de prise en charge en pratique courante.

Je vous remercie pour votre attention.

M^{me} PARIS. – Merci beaucoup pour cette présentation claire d'un dossier que j'ai trouvé assez complexe. Les courageux rapporteurs de ce dossier sont Jacques et Pauline. Je ne sais pas qui souhaite parler en premier. Jacques a l'air prêt. C'est vous qui décidez, je vous laisse décider.

M. ORVAIN. – Effectivement, c'est bien résumé. C'est un dossier qui est assez complexe techniquement. Nous avons eu beaucoup d'échanges qui nous ont permis de clarifier un peu le sujet. Et surtout, ce qui est intéressant, c'est que dans le rapport, il y a beaucoup de schémas qui aident à bien comprendre, en particulier dans les comparaisons, ce qui est comparé ? Quel produit est comparé à quel autre produit ? Cela permet d'y voir un peu plus clair.

Pour le résumer, je dirais que comme cela a été dit, que le gros problème que l'on a, c'est à quoi on compare Calquence ? Dans l'essai pivot qui a été présenté, on a l'impression que c'est un essai de phase II qui essaye de voir s'il faut le mettre en association ou pas, et qui prend comme comparaison le chlorambucil, ce n'est peut-être pas un comparateur historique, mais ce n'est tout de même pas le nec plus ultra de la pharmacopée dans ce domaine. Nous voyons bien qu'il nous manque une comparaison avec un traitement plus contemporain, en particulier le venetoclax. Nous n'avons pas cette comparaison.

L'industriel va recourir à une méta-analyse en réseau, avec beaucoup d'outils de traitement, et nous voyons, comme le chef de projet l'a montré, que la comparaison est intéressante. Elle est faite avec le recours d'une méta-analyse très longue, c'est-à-dire que chacun des produits comparés est à un bout du réseau. Cela veut dire qu'il y a sans doute beaucoup d'incertitudes qui se cumulent aux incertitudes, et l'on ne sait pas au bout du compte, si l'on peut faire crédit à la comparaison qui sort de cette méta-analyse en réseau.

En plus, par ailleurs, l'un des problèmes majeurs, c'est que nous n'avons pas de proportionnalité des hazard-ratios. Ce qui veut dire que potentiellement, on a des produits qui n'agissent pas exactement de la même façon au cours de la maladie. On va faire des comparaisons multiples et on a dérivé des différentiels d'activité sur des choses qui ne sont peut-être pas exactement du même ordre en termes de physiopathologie.

Avant de présenter le dossier en groupe technique, nous avons conclu avec les chefs de projet que cela constitue une réserve majeure. C'est-à-dire que nous n'avons pas vraiment la comparaison ou les données qui nous permettraient de conclure sur l'efficacité. Nous avons

estimé que c'était une réserve majeure, à la fois par le fait que nous n'avions pas de comparaison directe avec un comparateur de référence aujourd'hui, et par le fait que sur la méta-analyse en réseau qui était produite derrière, il y avait beaucoup d'incertitudes liées à ces absences de proportionnalité des hazard-ratios. On avait donc conclu à la réserve majeure.

Nous avons repassé cela dans le groupe technique. La discussion a plutôt orienté plutôt que réserve majeure, une incertitude majeure. Cela nous orientait plutôt là-dessus, en sachant que la décision n'était pas toujours facile.

Si j'ai bien compris, au niveau du bureau, on est revenu sur plutôt une réserve majeure. Nous en avons rediscuté avec les chefs de projet, effectivement, on penchait plutôt pour une réserve majeure, en considérant que nous n'avions pas les données qui permettaient de conclure sur cette efficience.

C'est ce qui a été présenté aujourd'hui. Nous en avons beaucoup discuté. Nous sommes restés sur cette proposition de réserves majeures.

M^{me} PARIS. – Merci beaucoup. Pauline.

M^{me} CHAUVIN. –Je voulais remercier vraiment le chef de projet pour son aide, pour la compréhension du dossier, parce qu'effectivement, c'est un dossier très compliqué. Vous avez vu à travers son exposé, ses qualités de pédagogue pour réussir à vraiment rentrer dans le dossier et comprendre certains points qui sont vraiment très techniques.

Je vais être assez courte parce que Jacques a bien résumé le principal gros problème que l'on a dans ce dossier, ce sont les données. Pour la mesure de l'efficience, les paramètres utilisés sont basés sur une méta-analyse en réseau que l'on considère comme pas valide, à cause de ces problèmes, entre autres, d'hypothèses des risques proportionnels, qui ne sont pas vérifiés. Les industriels ont essayé de proposer d'autres approches, mais cela n'était pas suffisant pour contourner le problème, et pour arriver à des estimations, des faits relatifs de survie globale, qui permettaient de renseigner l'efficience. Voilà pourquoi l'on propose ces réserves majeures.

L'autre point sur lequel je voulais revenir, mais vu que l'on propose une réserve majeure, cela ne changera pas grand-chose. Nous avons encore un dossier où l'on a une différence entre le périmètre de l'avis d'efficience et le périmètre de remboursement. Dans tous les cas, nous n'aurions pas pu fournir une information sur l'efficience sur l'ensemble du périmètre de remboursement.

Ce qu'il faut savoir, c'est qu'en gros, la population d'analyse de l'avis d'efficience représente à peu près 42 % de la population du périmètre de remboursement. C'est aussi à regretter dans cet avis, par rapport aux informations que nous aurions espéré fournir au CEPS.

C'est à peu près tout ce que je voulais dire sur ce dossier très compliqué. Je remercie les chefs de projet pour tous les échanges que nous avons eus dessus.

M^{me} PARIS. – Merci beaucoup. Y a-t-il d'autres commentaires sur ce dossier ? Je ne vois pas de main levée, sauf erreur de ma part. Devrait-on le voter ou avez-vous besoin de quelques minutes de réflexion ? C'est rare que nous n'ayons aucun commentaire. Il y a eu beaucoup de discussions en groupe technique. Peut-être que tout le monde a épuisé ses arguments et que maintenant, nous sommes prêts à converger.

Est-ce que l'on a le quorum pour voter, Samantha ?

M^{me} FERNANDES, HAS. – Oui, tout à fait.

M^{me} PARIS. – Catherine.

M^{me} LE GALES. – Merci beaucoup. On a effectivement beaucoup discuté sur ce dossier. Il a été extrêmement bien analysé et présenté pour qu'on en comprenne totalement les enjeux, à la fois au plan écrit et sur le plan de la présentation de ce matin. C'est toujours regrettable de terminer par trois réserves majeures, mais c'est comme ça.

La question que je me posais est plus de forme. Ma remarque porte sur la conclusion que j'ai trouvée assez courte, puisqu'elle fait trois lignes ou deux lignes et demie. Je me demandais si vraiment on n'avait rien. On n'a vraiment pas d'autres informations que l'on souhaite partager sur ce dossier ? C'est assez rare qu'on n'ait vraiment pas d'autres considérations à faire. Mais cela peut refléter évidemment le fait que dans ce dossier, il y a beaucoup de faiblesses méthodologiques, et qu'évidemment, on ne peut pas aller au-delà.

Mon deuxième point est plus un point de détail.

M^{me} PARIS. – Catherine.

M^{me} LE GALES. – Vous parlez de « profondes sources d'incertitude », je me suis demandé pourquoi vous aviez utilisé ce qualificatif. Parce que comme les mots sont très importants et le choix des mots est important. Dans un paragraphe, vous parlez de la conformité méthodologique et vous dites : « bien que non conclusive, l'évaluation économique permet de mettre en évidence de multiples et profondes sources d'incertitudes. » Je me suis interrogée. Je ne suis pas du tout opposée à l'idée d'utiliser cet adjectif, mais je me demandais ce qui vous avait motivés. Mais peut-être que l'on peut garder cette discussion pour plus tard. Ce n'est peut-être pas tout à fait le moment. Je voulais juste faire cette remarque sur le choix de cet adjectif pour qualifier les sources d'incertitude dans le cas particulier de ce dossier. De toute façon, cela ne change pas rien à la conclusion vers laquelle on se dirige. Merci.

Un chef de projet, pour la HAS.- Je veux bien répondre sur le terme, parce que c'est moi qui l'ai écrit. Je ne saurais pas le justifier de façon extrêmement précise, mais il fait tout de même référence au fait que l'incertitude est très forte et que l'on ne voit pas de piste à court ou moyen terme pour la lever. Parfois, on a l'impression qu'il y a une incertitude qui pourrait être levée par la production de certaines données. C'est vrai que là, on a l'impression d'être très loin de pouvoir la lever.

C'est le cas sur de multiples aspects de l'évaluation, à la fois sur l'extrapolation où l'on a vu que les fonctions d'extrapolation divergeaient assez fort et assez vite, et sur la comparaison entre les produits, la comparaison indirecte où l'on voit que l'industriel a cherché à approfondir précisément les approches pour essayer de comparer les produits. Il se trouve qu'une des approches qui était éventuellement prometteuse n'a pas été menée au bout. Mais nous avons aussi eu le sentiment qu'elle aurait peut-être permis de lever la réserve, mais certainement pas l'incertitude.

M. ORVAIN. – Là, tu fais référence aux survies moyennes qui auraient pu être utilisées, mais cela ne résout pas le problème que l'on peut avoir des produits qu'ils vont avoir une meilleure efficacité en début de traitement et un bénéfice qui s'altère après. Effectivement, la moyenne permet de gommer tout ça, parce que cela met des plus et des moins, et qu'à la fin on a une moyenne, mais cela n'exclut pas le problème. Ce n'est pas tout à fait la même chose d'avoir un traitement qui va être très efficace en début, et moins après, par rapport à un autre qui va être plus efficace sur le long terme.

L'artifice statistique permet effectivement, si l'on est d'accord, de résoudre le problème d'un point de vue statistique, mais d'un point de vue de la qualité de vie et de ce que vivent les patients, ce n'est pas du tout la même chose. Pour moi, cela n'aurait pas résolu le problème de fond qui est de qualifier la qualité de vie des patients par ces (*inaudible 0.13.31 enregistrement 0004*), cela reste une profonde incertitude.

M^{me} PARIS. – Vos questions me conduisent à poser une question au chef de projet. Je ne sais pas si tu seras en mesure d'y répondre immédiatement d'ailleurs. Mais ces incertitudes et ces problèmes de comparaison qui nous ont tellement gênés ont-ils une conséquence en termes d'évaluation par la CT. C'est-à-dire, est-ce que le niveau d'ASMR n'était pas celui revendiqué ? Y a-t-il une demande de données complémentaires ? C'est juste une question.

Un chef de projet, pour la HAS.- Je ne suis pas sûr que la CT ait encore validé définitivement l'avis.

M^{me} PARIS. – D'accord. Entendu.

M. ORVAIN. – Je suis d'accord avec Valérie que ce serait intéressant de savoir quel peut être l'avis de la CT, parce que nous voyons bien que c'est lié. C'est-à-dire que ce dont on parle, c'est l'efficacité.

M^{me} PARIS. – Tout à fait, une des choses qui nous a beaucoup gênés, c'est bien cette question clinique, l'efficacité comparative.

Un chef de projet, pour la HAS.- Oui, et il y a eu beaucoup de produits qui sont arrivés. On le voit par le nombre d'avis que nous voyons passer dans cette indication. Il y a eu beaucoup de produits récemment.

M. ORVAIN. – Je trouve dommage que sur des dossiers comme ceux-là, il n'y ait pas plus d'interactions avec la CT. Les avis sont tout de même à peu près sur les mêmes données.

M^{me} PARIS. – De quoi parles-tu ?

M. ORVAIN. – Un peu moins à cause du confinement (*inaudible 0.15.22 enregistrement 0004*) sur la cohérence des avis des deux commissions.

M^{me} PARIS. – Logiquement, ce qui gêne l'une des commissions devrait gêner l'autre ou en partie.

A propos de la conclusion en trois lignes, cela ne m'avait pas frappé. C'est vrai que les conclusions plus argumentées sont développées au-dessus. Du coup, je me demande ce que l'on fait d'habitude. Et pourquoi cela a-t-il frappé Catherine ? D'habitude, le répète-t-on ? Je pense qu'il n'y a pas de problème de fond ? C'est plutôt un problème de forme. Tu as l'impression que c'est cohérent ? C'est vrai que l'argumentaire est plus haut. Je ne sais pas s'il y a un intérêt à le répéter ou non. Catherine cela t'a frappé par rapport aux pratiques habituelles ?

M^{me} LE GALES. – Cela m'a frappée parce que parfois, on a la possibilité de donner quelques informations, y compris quantitatives, partielles, etc., où des points un peu plus généraux sont mis en exergue. Là, c'est vrai que malheureusement, la qualité de ce dossier, et tout ce qui a été discuté, fait que l'on voit bien que l'on a beaucoup de difficultés à pointer d'autres choses que cette conclusion.

Cela m'a juste un tout petit peu, non pas étonnée, mais par exemple avant, on parle de l'impact budgétaire, etc. C'est vrai que souvent, on a des conclusions qui sont un peu plus longues. Mais effectivement, dans le cadre de ce dossier, malheureusement, il y a très peu d'informations économiques ou générales que l'on peut mettre en exergue. De toute façon, les informations importantes sont juste avant, donc il n'y a pas de souci. Mais c'est vrai que c'est assez rare que l'on n'arrive pas à sortir quelque chose pour parler un peu crûment.

M^{me} PARIS. – C'est souvent une redite de ce qui a été dit au-dessus. Après, peut-être que c'est nécessaire.

M. ORVAIN. – Je ne dirais pas que c'est un mauvais dossier dans le sens où ils ont utilisé des tas d'approches. Techniquement, le dossier est assez sophistiqué. Mais le problème, c'est que l'on arrive à une forme d'équilibrisme. C'est-à-dire : Bravo l'artiste, tu arrives à tenir sur l'équilibre, mais on ne sait pas si tu ne vas pas te casser la figure dans le pas suivant. Ce n'est pas un mauvais dossier techniquement.

M^{me} LE GALES. – Parfaitement, je ne devrais pas utiliser ce mot, mais je pense simplement que toutes les difficultés méthodologiques, même si c'est techniquement, c'est très bien au plan méthodologique, ils ont fait avec ce qu'ils avaient, mais il n'en demeure pas moins que nous ne sommes pas en situation de mettre en lumière des informations économiques qui sont absolument indispensables pour pouvoir éclairer la décision publique. Malheureusement, c'est comme ça. Alors que dans certains dossiers, on arrive à trouver des points partiels qui paraissent utiles et importants.

Je trouve que ce qui est assez intéressant, c'est comme vous le dites, c'est que c'est une indication pour laquelle on a beaucoup de nouveaux produits qui se sont présentés sur le marché récemment. Là évidemment pour positionner celui-ci par rapport au dernier, cela va être extrêmement compliqué. On ne sait pas le faire, on ne peut pas le faire.

M^{me} PARIS. – J'imagine que c'est une situation qui se présente dans d'autres cas. Lorsque l'on a énormément d'arrivées dans une même indication, c'est compliqué pour les essais cliniques et pour l'évaluation d'avoir tous les comparateurs pertinents.

Je vous propose que l'on passe au vote. Carine, tu veux appeler.

Mme FERRETTI procède à l'appel nominatif.

Résultat du vote :

Favorable : 20 voix

(L'avis est adopté.)

M^{me} PARIS. – Merci encore.