



HAUTE AUTORITÉ DE SANTÉ

AVIS D'EFFICIENCE

KYPROLIS[®] (carfilzomib)
AMGEN

Date de validation par la CEESP : 10 mai 2016

Le présent avis est publié sous réserve des droits de propriété intellectuelle

Ce document a été validé par la Commission Evaluation économique et de santé publique
en mai 2016

© Haute Autorité de santé – 2016

HAS / Service évaluation économique et santé publique

Sommaire

Abréviations	4
1. Avis de la Commission Evaluation Economique et de Santé Publique	6
1.1 Objectif et contexte de l'étude	6
1.2 Conformité de l'étude médico-économique aux recommandations méthodologiques de la HAS.....	6
1.2.1 Analyse coût-efficacité.....	6
1.2.2 Analyse d'impact budgétaire	6
1.3 Conclusion de la CEESP sur l'efficience.....	6
1.4 Données complémentaires.....	8
2. Synthèse de l'analyse critique	10
3. Annexe 1 – Contexte de la demande	11
3.1 Objet de la demande	11
3.2 Produit et indication concernés par la demande	11
3.3 Historique d'autorisation de mise sur le marché	12
3.4 Historique du remboursement	13
4. Annexe 2 - Analyse critique détaillée de l'étude médico-économique	14
4.1 Documents support de l'analyse critique	14
4.2 Objectif de l'étude médico-économique proposée.....	14
4.2.1 Objectif tel que proposé par les auteurs	14
4.2.2 Analyse critique de l'objectif	14
4.3 Choix structurants concernant l'étude économique.....	14
4.3.1 Les choix structurants tels que présentés par les auteurs	14
4.3.2 Analyse critique concernant les choix structurants.....	16
4.4 La modélisation	17
4.4.1 La modélisation telle que présentée par les auteurs	17
4.4.2 Analyse critique concernant la modélisation	27
4.5 Mesure et valorisation des états de santé.....	31
4.5.1 Evaluation des résultats de santé telle que présentée par les auteurs	31
4.5.2 Analyse critique de l'estimation des résultats de santé	32
4.6 Mesure et valorisation des coûts.....	33
4.6.1 Evaluation des coûts telle que présentée par les auteurs	33
4.6.2 Analyse critique de l'évaluation des coûts.....	39
4.7 Présentation des résultats et analyses de sensibilité	39
4.7.1 Présentation par les auteurs.....	39
4.7.2 Analyse critique de la présentation des résultats et de l'analyse de sensibilité	46
Annexe 3 – Faisabilité de la méta-analyse	47
Annexe 4 – Echange avec l'industriel	50
Bibliographie	57

Abréviations

ACSH	Autogreffe de cellules souches hématopoïétiques
AIC	Critères d'informations d'Akaike
ALD	Affection longue durée
AMM.....	Autorisation de mise sur le marché
ASMR....	Amélioration du service médical rendu
ATU	Autorisation temporaire d'utilisation
AVG	Année de vie gagnée
BIC	Critère d'information bayésien
BSC.....	Best Supportive Care – Meilleurs soins de soutien
CEACs ..	Courbes d'acceptabilité de coût-efficacité
CEESP..	Commission évaluation économique et de santé publique
CEPS	Comité économique des produits de santé
CHMP ...	Committee for Medicinal Products for Human Use
CSH	Cellules souches hématopoïétiques
DLP	Doxorubicine liposomale pégylée
ECOG ...	Eastern Cooperative Oncology Group
EGB	Echantillon généraliste de bénéficiaires
EI.....	Effet Indésirable
EMA	European Medicines Agency
ENCC....	Etude Nationale Commune de Coûts
EORTC .	European organization for research and treatment of cancer
EQ-5D ...	EuroQoL à 5 dimensions
ERd	EMPLICITI® + REVLIMID® + dexaméthasone faible dose
GHM.....	Groupe Homogène de Malades
GHS	Groupe Homogène de Séjour
HAS.....	Haute Autorité de Santé
HR.....	Hazard ratio
HT.....	Hors taxe
HTA.....	Health technology assessment
IC	Intervalles de confiance
IMWG....	l'International Myeloma Working Group
ITT.....	Intention de traiter
IV.....	Intra veineuse
KRd	KYPROLIS® + REVLIMID® + dexaméthasone faible dose
MCO.....	Médecine Chirurgie Obstétrique
NABM....	Nomenclature des actes de biologie médicale
NGAP	Nomenclature Générale des Actes Professionnels
NICE	National Institute for Health and Care Excellence
NFS.....	Numération Formule Sanguine
NR.....	Non renseigné

ORR Overall Response Rate
OS / SG. Overall Survival – Survie globale
PanoVd. FARYDAK® + VELCADE® + dexamethasone faible dose
PPTTC .. Prix Public Toutes Taxes Comprises
PFHT Prix Fabricant Hors Taxe
PFS / SSP Progression-Free Survival – Survie sans progression
PMSI Programme de Médicalisation des Systèmes d'Information
PPTTC .. Prix Public Toute Taxe Comprise
PSA Probabilistic Sensitivity Analysis – Analyse de sensibilité probabiliste
QALY Quality-adjusted life year
QdV Qualité de vie
RC Réponse complète
RCP Réunion de Concertation Pluridisciplinaire
Rd Revlimid® + dexamethasone faible dose
RDCR.... Ratio différentiel coût-résultat
SC Sous-cutané
SMR Service Médical Rendu
TTC Toutes Taxes Comprises
TTD Temps avant arrêt de traitement

1. Avis de la Commission Evaluation Economique et de Santé Publique

1.1 Objectif et contexte de l'étude

L'étude présentée par l'industriel évalue l'efficience de Kyprolis® (carfilzomib) en association avec Revlimid® (légalidomide) et la dexaméthasone dans le traitement du myélome multiple, chez les patients adultes qui ont reçu au moins un traitement antérieur.

Cette évaluation soutient une demande d'inscription du laboratoire AMGEN sur la liste des médicaments agréés à l'usage des collectivités et divers services publics, pour laquelle l'industriel revendique une ASMR de niveau II.

Le chiffre d'affaires prévisionnel TTC la 2^{ème} année pleine de commercialisation de Kyprolis® est attendu être supérieur à 20 millions d'euros PPTTC.

1.2 Conformité de l'étude médico-économique aux recommandations méthodologiques de la HAS

1.2.1 Analyse coût-efficacité

La méthode sur laquelle repose l'évaluation médico-économique relative au produit est acceptable, bien qu'elle soulève des réserves importantes.

Ces réserves sont détaillées dans l'annexe et présentées dans le tableau de synthèse page 9.

1.2.2 Analyse d'impact budgétaire

Le choix d'intégrer dans le dossier une analyse d'impact budgétaire est laissé à la libre appréciation de l'industriel. Dans le cadre de ce dossier, aucune analyse d'impact budgétaire n'a été fournie par l'industriel.

1.3 Conclusion de la CEESP sur l'efficience

Contexte de l'évaluation et portée de la conclusion

La CEESP souligne que les données cliniques disponibles dans le myélome multiple n'ont pas permis d'évaluer l'efficience de Kyprolis® (carfilzomib) par rapport à l'ensemble de ses comparateurs dans la stratégie thérapeutique (méta-analyse non faisable), mais uniquement versus les associations de légalidomide et dexaméthasone (Rd) et de bortezomib et dexaméthasone (Vd).

La CEESP souligne également que l'évaluation présentée ne peut être interprétée que dans le contexte actuel de la prise en charge. L'arrivée prochaine de nombreuses options thérapeutiques dans le traitement du myélome multiple rendra caducs les résultats présentés dans cet avis.

Par ailleurs, des essais sont en cours, qui laissent présager une extension d'indication de carfilzomib à la première ligne du myélome multiple dans les années à venir.

La CEESP attire l'attention sur le risque que le prix auquel aboutira la négociation sur cette indication restreinte pourrait avoir un effet déterminant, d'une part, sur le prix du carfilzomib dans ses futures autres indications et, d'autre part, sur le prix d'autres molécules en cours d'évaluation dans le myélome multiple.

Résultats de l'évaluation

D'après les données et les hypothèses fournies par l'industriel, Kyprolis® (carfilzomib), en association avec Revlimid® (légalidomide) et la dexaméthasone (KRd), est associé à un RDCR versus légalidomide et la dexaméthasone (Rd) de 287 035 €/QALY, 292 960 €/QALY et 292 289 €/QALY pour la 2^{ème}, 3^{ème} et 4^{ème} ligne respectivement. L'association bortezomib et dexaméthasone (Vd) est dominée en 2^{ème} et 3^{ème} ligne.

L'incertitude générée par les données entrées dans le modèle, notamment l'efficacité de Kyprolis® sur la survie, est importante. Sous les hypothèses et les choix méthodologiques retenus dans l'analyse de référence, le choix de l'association KRd par rapport à Rd est efficace avec un degré de certitude de 80% pour une disposition à payer de 390 000 €/QALY (2^{ème}, 3^{ème} et 4^{ème}).

L'analyse critique de la méthodologie a permis de mettre en évidence plusieurs éléments méthodologiques qui concourent à rendre l'estimation du RDCR incertaine.

- La modélisation de la survie (SG et SSP) dans le bras de référence Rd sur la base d'une population de registre recevant un traitement au choix du médecin (TCM) conduit à la sous-estimation du RDCR.
- L'application d'un $HR_{KRdvsRd}$ estimé sur une population d'essai clinique avec un meilleur pronostic que la population ayant servi à modéliser la survie sous Rd conduit à la sous-estimation du RDCR, en raison d'une potentielle surestimation de l'effet traitement de KRd vs Rd.

Cela est d'autant plus préoccupant que le $RDCR_{KRdvsRd}$ est très sensible à l'estimation de l'effet traitement de KRd vs Rd sur la survie globale issue de l'essai ASPIRE, dans lequel la médiane de survie n'est pas atteinte lors de l'analyse intermédiaire prévue. En appliquant la variabilité du hazard ratio correspondant à l'intervalle de confiance observé dans l'essai de phase III ASPIRE pour la survie globale ($HR=0,74$; $IC_{95\%}$ 0,59 – 0,94), le RDCR varie entre 213 000 €/QALY et 538 000 €/QALY pour la sous-population en première rechute. L'incertitude est encore plus importante pour les patients en 2^{ème} (206 000 € – 673 000 €) et 3^{ème} rechutes (196 000 € – 805 000 €).

- Les méthodes mobilisées pour valoriser la qualité de vie conduisent à appliquer des valeurs d'utilité avant progression vraisemblablement surestimées, ce qui est favorable à Kyprolis®.

En conclusion, les choix de modélisation tendent globalement à sous-estimer le RDCR de Kyprolis® versus Rd.

L'efficience en vie réelle de Kyprolis®, en association avec Revlimid® et la dexaméthasone, dépendra majoritairement de la durée de traitement et de la durée de l'effet traitement dans le temps.

Le RDCR estimé considère un effet traitement de KRd vs Rd sur une durée 36 mois. Or, l'industriel présente plusieurs arguments en faveur d'une hypothèse d'effet traitement qui serait maintenu au fil du temps. Dans ce cas, l'efficience de Kyprolis® serait améliorée en pratique réelle. Cependant, la CEESP souligne que, même dans cette situation optimiste, le $RDCR_{KRd vs Rd}$ reste extrêmement élevé, puisqu'il est compris entre 183 000 €/QALY et 208 000 €/QALY selon la ligne d'instauration.

Par ailleurs, si la durée de traitement par Kyprolis® est allongée, par exemple jusqu'à progression de la pathologie (posologie possible dans le RCP), les gains de survie obtenus par la persistance d'un effet traitement seraient compensés par l'augmentation des coûts de traitement.

Concernant les coûts de traitement, la CEESP note que l'impact d'une variation du prix de Kyprolis® est limité en raison de son association avec Rd, dont le prix facial est déjà élevé. Lorsque le prix de Kyprolis® est réduit de 50%, le RDCR n'est réduit que de 28% à 29%, selon la sous-population.

Compte tenu du coût associé à l'association Rd, sa durée de traitement aura un impact très important sur le RDCR de KRd en vie réelle. Or la durée moyenne de traitement de Rd est

largement augmentée avec l'adjonction de Kyprolis®, en raison d'une durée de survie sans progression plus longue. Cette analyse met en évidence deux problématiques associées à la négociation du prix de produits prescrits en association avec d'autres molécules déjà sur le marché. Premièrement, le rapport coût-résultat de Kyprolis® en adjonction à Rd est conditionné par le prix de Revlimid®, dont l'AMM a été obtenue en 2007, et la baisse du prix du seul Kyprolis® ne permettra pas de réunir les conditions de son efficience. Deuxièmement, l'adjonction de Kyprolis® va avoir un impact sur les volumes de vente de Revlimid® en augmentant les durées de traitements (ex : le lenalidomide est administré pendant 33 mois dans le bras Krd versus 22 mois dans le bras Rd).

En conclusion, la CEESP estime que le RDCR est extrêmement élevé. D'une part, la prise en charge du myélome multiple a bénéficié de progrès thérapeutiques récents, ce qui limite le bénéfice marginal apporté par l'adjonction de Kyprolis® à Rd. D'autre part, Kyprolis® est un traitement, dont la place dans la stratégie thérapeutique se définit en adjonction à un produit innovant (Revlimid®), dont le prix est déjà élevé (les coûts de traitement par Rd représentent plus de la moitié des coûts de KRd).

Dans l'interprétation de l'efficience de Kyprolis®, la CEESP insiste sur l'importance de prendre en compte les points suivants.

- Les données cliniques disponibles dans le myélome multiple n'ont pas permis d'évaluer l'efficience de Kyprolis® par rapport à l'ensemble de ses comparateurs. Par ailleurs, l'arrivée prochaine de nombreuses options thérapeutiques rendra caducs les résultats présentés dans un délai très proche.
- Le RDCR de Kyprolis® (carfilzomib), en association avec Revlimid® (lénalidomide) et la dexaméthasone (KRd) est estimé entre 287 000 €/QALY et 293 000 €/QALY vs Rd, en fonction de la ligne d'instauration de traitement (2^{ième} à 4^{ième} ligne).
- La CEESP souligne le caractère extrêmement élevé du RDCR estimé.
- La valeur du RDCR estimé est sensible à l'effet traitement de KRd vs Rd sur la survie globale, dont l'estimation est incertaine (médiane de survie non atteinte dans l'essai clinique ASPIRE), et plusieurs choix méthodologiques concourent vraisemblablement à sous-estimer le RDCR présenté.
- La CEESP souligne que l'efficience de Kyprolis® est aussi conditionnée par une potentielle renégociation du prix facial de Revlimid®, dont une augmentation du volume des ventes peut être anticipée. Lorsque le prix de Kyprolis® est réduit de 50%, le RDCR n'est réduit que de 28% à 29%, selon la sous-population.
- La CEESP souligne les enjeux en termes d'impact budgétaire soulevés par une éventuelle extension de l'utilisation de Kyprolis® en première ligne.

1.4 Données complémentaires

La CEESP recommande que l'évaluation soit actualisée dans un délai permettant, d'une part, d'intégrer les comparateurs de KRd arrivant prochainement sur le marché et, d'autre part, de réduire l'incertitude sur les données de survie globale de KRd et Rd.

Il reste qu'un travail exploratoire, réalisé par l'industriel à la demande de la HAS, a démontré l'impossibilité d'alimenter une évaluation économique incluant tous les comparateurs de KRd à

partir d'une méta-analyse en réseau conventionnelle à partir des essais disponibles au moment du dépôt de dossier. La CEESP souligne les efforts de l'industriel et l'intérêt de poursuivre un travail de veille sur les essais, dans la perspective de pouvoir à terme relier les différents réseaux et permettre une véritable évaluation médico-économique des traitements dans le myélome multiple.

2. Synthèse de l'analyse critique

Les points de critique identifiés dans l'analyse détaillée sont hiérarchisés selon trois niveaux¹.

- Réserve mineure (-) : élément non conforme aux recommandations en vigueur, mais qui est justifié ou dont l'impact attendu sur les conclusions est négligeable.
- Réserve importante (+) : élément non conforme aux recommandations en vigueur, avec un impact attendu important sur les conclusions (en particulier en termes d'incertitude).
- Réserve majeure (++) : élément non conforme aux recommandations en vigueur qui invalide tout ou partie de l'étude économique.

Tableau 1 : Synthèse de l'analyse critique de l'évaluation économique

Libellé de la réserve	-	+	++
Choix structurants			
Absence justifiée de plusieurs comparateurs : données cliniques disponibles incompatibles pour une méta-analyse en réseau.	-		
Modélisation			
Extrapolation des données Rd sur un nombre limité de fonctions paramétriques	-		
Sous-estimation potentielle de l'effet Vd (HR appliqué sur 18 mois contre 36 mois pour KRd)			
Estimation des courbes de survie			
- Courbes de survie (Rd) estimées sur des données de registre non représentatives du traitement évalué (Rd)	-		
- Populations différentes pour l'estimation de Rd et pour l'estimation de KRd (surestimation potentielle de l'effet du produit)		+	
- Hypothèses d'effet traitement (KRd, Vd) identique par ligne insuffisamment justifiées	-		
- Recours justifié à une analyse statistique ajustée (« Matched Adjusted Indirect Comparaison) pour intégrer Vd, dont le niveau de preuve est faible.	-		
Sous-estimation de la durée de traitement sous Kyprolis® (18 cycles) par rapport à ce qui peut être attendu en vie réelle (le RCP indique une possibilité de poursuivre Kyprolis® jusqu'à progression)	-		
Non modélisation du risque de troubles cardiaques pour Krd	-		
Mesure et valorisation des états de santé			
Recours au mapping avec choix d'un algorithme favorable au produit	-		
Méthode d'estimation de l'utilité par ligne de traitement favorable au produit	-		
Résultats et analyses de sensibilité			
Absence des analyses de sensibilité suivantes : - HT inférieur à 15 ans - sur l'hypothèse d'équivalence des HR sur les lignes de traitement - sur les HR non-ajustée de l'essai ASPIRE - sur une durée d'effet de Vd identique à KRd		+	

3. Annexe 1 – Contexte de la demande

3.1 Objet de la demande

Le dossier d'évaluation économique de KYPROLIS® est déposé auprès du CEESP par le laboratoire AMGEN dans le cadre d'une primo-inscription sur la liste des spécialités agréées à l'usage des collectivités et divers services publics.

La demande entre dans le cadre du décret du 2 octobre n°2012-116 :

- l'industriel revendique une ASMR II pour son indication ;
- le chiffre d'affaires annuel attendu a été estimé à [REDACTED] PPTTC après 2 années de commercialisation.

3.2 Produit et indication concernés par la demande

KYPROLIS® (carfilzomib) est un inhibiteur de protéasome de 2^{ème} génération qui empêche la dégradation de certaines protéines intra cellulaires et affecte de multiples cascades de signaux à l'intérieur de la cellule, entraînant *in fine* la mort de la cellule cancéreuse.

Indications concernées par la demande

KYPROLIS® en association avec REVLIMID® (légalidomide) et la dexaméthasone est indiqué dans le traitement du myélome multiple, chez les patients adultes qui ont reçu au moins un traitement antérieur.

Le myélome multiple (MM) est une hémopathie maligne caractérisée par la prolifération et l'accumulation de plasmocytes tumoraux, monoclonaux (même chaîne lourde et/ou légère) dans la moelle osseuse. Il s'agit d'une maladie rare : environ 1 % de tous les cancers (Palumbo et al. NEJM, 2011), on estime à 5 000 le nombre de nouveaux cas en 2012 en France (Monnereau et al. InVS, 2013). En dépit des progrès thérapeutiques récents tendant à la chronicisation de la maladie, le MM demeure incurable, les patients initialement répondeurs aux traitements disponibles développant progressivement des résistances à ces traitements.

Dans l'étude des registres d'hémopathie maligne de la Côte d'Or, de la Gironde et du Calvados faite par AMGEN, la durée médiane sans progression observée est d'environ 25 mois en première ligne, 18 mois en deuxième ligne, 12 mois en troisième ligne et 6 mois en quatrième ligne.

Stratégie thérapeutique

Chez les patients éligibles de moins de 65 ans, le traitement recommandé en première ligne est une chimiothérapie intensive associée à une autogreffe de cellules souches du sang périphérique (CSP).

Il n'existe pas de traitement standard pour une rechute ou progression du myélome multiple selon la Société Française d'Hématologie. La décision thérapeutique dépend de l'âge, des traitements antérieurs, de la durée de la première rémission et des circonstances de la rechute, de la disponibilité de CSP, de l'état général et des comorbidités (existence ou non d'une neuropathie ou d'une insuffisance rénale). De façon générale, si la durée de première réponse est importante, le traitement de première ligne pourra être à nouveau utilisé. Chez les jeunes patients, après traitement de rattrapage, une auto- ou allogreffe de cellules souches hématopoïétiques est proposée. Dans les autres cas, plusieurs associations sont utilisées : Velcade® (bortézomib) est utilisé (hors AMM) dans le cadre d'associations, notamment aux immuno-modulateurs (légalidomide ou thalidomide) et au cyclophosphamide. Le légalidomide (Revlimid®) est utilisé en association avec la dexaméthasone uniquement. Il est habituel d'administrer un immuno-modulateur chez un patient qui aura reçu du bortézomib en première ligne.

Pour les patients atteints d'un myélome ayant reçu au moins deux traitements antérieurs comprenant lénalidomide et bortézomib et dont la maladie a progressé pendant le dernier traitement :

- le pomalidomide (Imnovid®) en association à la dexaméthasone
- le panobinostat (Farydak®) est en cours d'évaluation par la Commission de la Transparence dans l'indication du myélome en rechute et/ou réfractaire ayant déjà reçu deux lignes de traitement antérieur.

Population cible

La population cible de Kyprolis® est représentée par l'ensemble des patients atteints d'un myélome multiple ayant reçu au moins un traitement antérieur (2^{ème} ligne et plus).

La prévalence du myélome multiple en France est d'environ 20 000 d'après la CNAMTS (répartition et taux des principales maladies des personnes en Affection Longue Durée au 31 décembre 2012 pour le régime général <http://www.ameli.fr>). Parmi eux environ 5 000 patients sont des cas incidents.

Considérant que le recours à un traitement de deuxième ligne concerne environ 60 %² des patients, la population cible incidente de Kyprolis® serait de l'ordre de 2 350 à 2 500 patients par an en 2^{ème} ligne. Au-delà de la deuxième ligne, compte tenu des nombreux critères de non inclusion de l'étude ASPIRE quant à l'utilisation antérieure d'inhibiteurs du protéasome notamment, la population cible n'est pas quantifiable.

La population prévalente revendiquée par l'industriel pour KRd est estimée entre 6 400 et 9 000 patients dont un peu moins de la moitié sont en deuxième ligne et l'autre moitié en 3^{ème} ligne et plus.

Essais en cours sur le produit

En plus des deux essais utilisés dans l'évaluation économique (ASPIRE³ et ENDEAVOR⁴), plusieurs essais de phase III sont en cours, dont deux en première ligne de traitement chez des patients nouvellement diagnostiqués.

- KRd compared with Velcade+Rd in patients with newly diagnosed myeloma [NCT01863550]
- Kyprolis®, melphalan, and prednisone compared with Velcade, melphalan, and prednisone in patients with newly diagnosed myeloma [NCT01818752]. Essai CLARION - analyse finale août 2019
- Kyprolis® and dexamethasone given on a once-weekly or a twice-weekly dosing schedule in patients with relapsed and refractory myeloma [NCT02412878]. Essai ARROW -analyse finale mars 2018.
- A Randomized, Open-label, Phase 3 Study of Carfilzomib vs Best Supportive Care in Subjects With Relapsed and Refractory Multiple Myeloma [NCT01302392]. Premiers résultats disponibles.

3.3 Historique d'autorisation de mise sur le marché

La spécialité Kyprolis® a obtenu, dans l'Union Européenne, la désignation de médicament orphelin en date du 3 juin 2008.

L'AMM européenne a été accordée le 19/11/2015.

² Kumar SK, Therneau TM, Gertz MA, et al. Clinical course of patients with relapsed multiple myeloma. Mayo Clin Proc 2004;79:867-874

³ Phase 3 Study Comparing Carfilzomib, Lenalidomide, and Dexamethasone (CRd) vs Lenalidomide and Dexamethasone (Rd) in Subjects With Relapsed Multiple Myeloma (NCT01080391)

⁴ Phase 3 Study With Carfilzomib and Dexamethasone Versus Bortezomib and Dexamethasone for Relapsed Multiple Myeloma Patients) (NCT01568866)

3.4 Historique du remboursement

Il s'agit de la première demande de remboursement de carfilzomib (Kyprolis®) en France. Cette demande ne concerne que le carfilzomib en association avec le légalidomide et la dexaméthasone.

Kyprolis® existe dans trois conditionnements : flacon de 60mg, flacon de 30mg et flacon de 10mg.

Le laboratoire revendique un PPTTC de [REDACTED] €/mg.

Selon l'industriel, la dépense annuelle moyenne par patient à ce prix est estimée à [REDACTED] € PPTTC.

Le montant remboursable à 5 ans, estimé par l'industriel sur la base d'une population rejointe de [REDACTED] patients est de [REDACTED] euros par an.

4. Annexe 2 - Analyse critique détaillée de l'étude médico-économique

4.1 Documents support de l'analyse critique

Quatre documents transmis lors du dépôt de dossier par l'industriel à la HAS ont fait l'objet d'une analyse critique :

- Un rapport de présentation ;
- Un rapport technique de l'étude d'efficience ;
- Une version électronique du modèle d'efficience au format Excel ;
- La bibliographie du rapport technique.

Une liste de questions techniques a été adressée à l'industriel. Une réponse écrite a été fournie et l'analyse critique tient compte de cette réponse ainsi que des 14 annexes jointes avec le dossier d'efficience.

Des documents complémentaires ont également été fournis dans le dossier :

- Le dossier technique de la CT.

L'analyse critique évalue la recevabilité de l'évaluation économique au regard du guide méthodologique en vigueur (HAS, 2011).

4.2 Objectif de l'étude médico-économique proposée

4.2.1 Objectif tel que proposé par les auteurs

L'objectif de l'étude déposée par AMGEN est l'évaluation, chez les patients adultes atteints d'un myélome multiple en rechute, du rapport différentiel coût-résultat de Kyprolis® (carfilzomib) en association avec le Revlimid® et la dexaméthasone (KRd) par ligne de traitement (2^{ème}, 3^{ème} ou 4^{ème} ligne de traitement) comparé aux traitements actuellement prescrits.

4.2.2 Analyse critique de l'objectif

L'objectif est clairement énoncé.

Il est clairement précisé que l'évaluation ne peut être interprétée que dans le contexte actuel de prise en charge. L'arrivée prochaine de plusieurs options thérapeutiques rendra caducs les résultats présentés dans un délai très proche.

4.3 Choix structurants concernant l'étude économique

4.3.1 Les choix structurants tels que présentés par les auteurs

► L'analyse économique et le choix du critère de résultat

L'analyse mise en œuvre est de type coût-utilité prenant le QALY comme critère de résultat de santé.

Une analyse de type coût-efficacité en coût par année de vie gagnée est également proposée.

► La perspective

La perspective est qualifiée de collective par l'industriel. Elle renvoie aux financeurs des soins (assurances maladies, patients).

► L'horizon temporel et l'actualisation

Le modèle met en œuvre un horizon temporel vie entière.

L'analyse de référence prévoit un taux d'actualisation à 4% pour les coûts et les effets de santé. Des taux d'actualisation de 0 % et 2,5 % ont été testés dans des analyses de sensibilité.

► La population d'analyse

La population d'analyse est définie par rapport à l'indication de l'AMM, et correspond aux patients adultes atteints de myélome multiple en rechute ayant reçu au moins 1 traitement antérieur.

Aucun RDCR ne peut être estimé dans cette population groupée, en raison de l'hétérogénéité des comparateurs selon la ligne de traitement.

L'analyse de référence est donc présentée en distinguant 3 sous-populations d'analyse, définies sur les patients atteints de myélome multiple en 2^{ème}, 3^{ème} et 4^{ème} lignes de traitement.

L'efficience de KRd dans les lignes plus avancées n'a pu être évaluée en raison (1) du nombre très faible de patients en 5^{ème} ligne de traitement dans le registre, (2) du fait que l'étude ASPIRE n'incluait aucun patients au-delà de la 4^{ème} ligne de traitement.

► Les stratégies comparées

Les comparateurs dans le modèle ont été sélectionnés sur la base de données observationnelles françaises d'utilisation (Étude RCP Bretagne ; registre des hémopathies malignes) et de la possibilité de mobiliser des données d'efficacité relative fiables sur la survie et la survie sans progression.

Les données d'utilisation indiquent que les traitements les plus utilisés en 2^{ème} ligne sont le Revlimid® (légalidomide) et le Velcade® (le plus fréquemment avec la dexaméthasone, mais d'autres associations peuvent être utilisées). A partir de la 3^{ème} ligne, Levact® ou Imnovid® deviennent des options plus fréquentes, mais Revlimid® et Velcade® restent majoritaires.

Des données sont disponibles pour Rd (essai ASPIRE) et Vd (essai ENDEAVOR). Ces essais ne peuvent être liés via une méta-analyse en réseau (faisabilité de la méta-analyse disponible en annexe 3). Les auteurs ont dû recourir à une analyse statistique groupée des données individuelles.

Les comparateurs pris en compte sont donc uniquement Rd et Vd (pour chaque ligne de traitement).

Le tableau ci-dessous reprend les traitements non pris en compte, ainsi que l'argument présenté par l'industriel.

Alternatives	Justification
Autogreffe	Comparateur non pertinent en R/R. Traitement standard chez les patients éligibles dont le myélome est nouvellement diagnostiqué.
Bortézomib en monothérapie	Non utilisé en pratique
Bortézomib en association à la doxorubine liposomale pégylée (DLP)	Non utilisé en pratique
Farydak+Velcade+dexaméthasone	AMM aout 2015 en 3 ^{ème} ligne et + après VELCADE® et un agent immunomodulateur. Pas de prix
Emplicit+Revlimid+dexaméthasone	AMM janvier 2016 en 2 ^{ème} ligne et + Pas de prix
Ixazomib+Revlimid+dexaméthasone	L'indication exacte est en cours d'évaluation auprès de l'EMA.
Kd	L'indication exacte est en cours d'évaluation auprès de l'EMA.

4.3.2 Analyse critique concernant les choix structurants

► L'analyse économique et le choix du critère de résultat

Le choix d'une analyse principale de type coût-utilité complétée par une analyse coût-efficacité prenant l'année de vie gagnée comme critère de résultat est conforme aux recommandations méthodologiques de la HAS.

► La perspective

Le guide HAS définit la perspective collective comme suffisamment large pour tenir compte de l'ensemble des parties prenantes concernées par les interventions, que ce soit en termes d'effets de santé ou de financement.

La perspective adoptée est qualifiée de collective sans lien véritable avec cette définition, puisqu'elle n'est justifiée qu'en référence aux coûts inclus dans l'analyse.

La perspective réellement retenue est une perspective « financeurs ». Ce choix est acceptable dans le cadre de ce dossier.

► L'horizon temporel et l'actualisation

Le choix d'un horizon temporel définit vie entière est cohérent avec la nature de la pathologie.

Le taux d'actualisation retenus dans l'analyse de référence est conforme aux recommandations méthodologiques de la HAS.

► La population d'analyse

La définition de la population d'analyse en trois sous-populations est conforme à l'hétérogénéité des stratégies de prise en charge selon la ligne de traitement, observées en vie réelle.

► Les comparateurs

Le myélome multiple connaît actuellement une phase d'innovation importante, avec l'arrivée de nombreuses nouvelles options thérapeutiques

La comparaison avec toutes les options pour lesquelles des données ont été publiées aurait été souhaitable. Le prix de ces options n'étant pas connu, des hypothèses auraient pu être testées.

Néanmoins, il n'est pas possible de mener une méta-analyse sans recourir à des hypothèses cliniques fortes (cf. Annexe 3), rendant impossible l'inclusion de comparateurs autres que Rd et Vd avec une méthode conventionnelle.

Alternatives	Acceptabilité de la justification
Bortézomibe en monothérapie	La non-utilisation en pratique est une justification insuffisante.
Bortézomibe en association à la doxorubine liposomale pégylée (DLP)	La non-utilisation en pratique est une justification insuffisante.
Farydak (panobinostat) +Velcade+dexaméthasone	Comparateur non pertinent en 2 ^{nde} ligne Comparateur pertinent en 3 ^{ème} ligne de traitement et plus Données disponibles Pas de prix au moment du dépôt de dossier
Emplicit (elotuzumab) +Revlimid+dexaméthasone	Comparateur pertinent en 2 ^{ème} ligne et + Données disponibles Pas de prix au moment du dépôt de dossier
Ninlaro (Ixazomib) +Revlimid+dexaméthasone	Approuvé par la FDA en novembre 2015 pour les patients ayant reçu au moins une ligne de traitement. Pas de prix au moment du dépôt de dossier
Chimiothérapie standard	Non discutée : option envisageable en 2 ^{nde} ligne chez les patients après rémission de longue durée.
Kd	Décision FDA d'extension d'indication en janvier 2016. En monothérapie pour les patients R/R ayant reçu au moins une ligne de traitement. Option intégrée dans des recommandations récentes (cf. NCCN Clinical Practice Guidelines in Oncology, version 2.2016).

4.4 La modélisation

4.4.1 La modélisation telle que présentée par les auteurs

► La population simulée

Trois sous-populations sont simulées, en fonction de la ligne d'instauration de KRd en 2^{ème} ligne, en 3^{ème} ligne et en 4^{ème} ligne.

Les trois sous-populations simulées dans le modèle sont issues du Registre des Hémopathies Malignes de Côte d'Or, Gironde et Calvados. Il s'agit de patients avec un myélome symptomatique en rechute.

Le registre inclut 148 patients en 2^{ème} ligne, 93 patients en 3^{ème} ligne et 49 patients en 4^{ème} ligne (Tableau 2).

Les auteurs comparent les caractéristiques des patients inclus dans le registre aux populations en intention de traiter (ITT) de l'essai ASPIRE (n = 792) et ENDEAVOR (n = 929), et soulignent que la population du registre est dans un état plus avancé de la pathologie que la population des essais cliniques ASPIRE et ENDEAVOR. Les patients du registre sont :

- plus âgés (médiane de 70 ans dans le registre versus 65 ans dans les essais cliniques), avec une clairance de la créatinine plus basse en moyenne (surtout versus l'essai ASPIRE) ;
- moins concernés par les greffes de cellules souches (33% à 39% dans le registre vs 56% à 58% dans les essais)
- avec un score ECOG, facteur de pronostic, inconnu pour plus de la moitié d'entre eux.

La distribution des traitements précédemment reçus est également hétérogène entre les différentes sources.

Tableau 2 : Caractéristiques de la population simulée et de celle des études clinique ASPIRE et ENDEAVOR

Caractéristiques	ASPIRE	ENDEAVOR	Registre d'hémopathies malignes régional 2 ^{ème} ligne	Registre d'hémopathies malignes régional 3 ^{ème} ligne	Registre d'hémopathies malignes régional 4 ^{ème} ligne
Nombre de patients (n)	792	929	148	93	49
BSA médiane	1.9	1.8			
Âge (ans)					
Médiane	64	65	69.5	70.6	70
Ecart	31 - 91	30 - 89	46 - 94	46 - 92	46 - 84
Age médian au diagnostic (années)			67	67	66
Distribution – Nombre de patients (%)					
18-64 ans	399 (50%)	433 (47%)	61 (41%)	36 (39%)	22 (45%)
Sexe					
Homme	447 (56%)	469 (51%)	78 (53%)	49 (53%)	28 (57%)
Score ISS					
	Au diagnostic	A la baseline	Au diagnostic	Au diagnostic	Au diagnostic
1	138 (17%)	417 (45%)	35 (24%)	21 (23%)	11 (22%)
2	193 (24%)	289 (31%)	32 (25%)	50 (34%)	35 (38%)
3	346 (44%)	223 (24%)	38 (29%)	32 (22%)	17 (18%)
Non su			31 (21%)	20 (22%)	10 (20%)
Statut de performance ECOG					
0 ou 1	717 (91%)	867 (93%)	43 (29%)	30 (32%)	17 (35%)
2+	75 (9%)	62 (7%)	7 (5%)	5 (5%)	3 (6%)
Non su			98 (66%)	58 (62%)	29 (59%)
% de patients greffés (antécédents)	446 (56%)	538 (58%)	49 (33%)	30 (32%)	19 (39%)
Distribution – Nombre de patients (%)					
1 ^{ère} ligne					
2 ^{ème} ligne	341 (43%)	464 (50%)			
3 ^{ème} ligne	259 (33%)	302 (33%)			
4 ^{ème} ligne et plus	192 (24%)	163 (18%)			
Traitement précédent					
Bortezomib	521 (66%)	502 (54%)	91 (61%)	83 (89%)	49 (100%)
Lenalidomide	157 (20%)	354 (38%)	7 (5%)	38 (41%)	37 (76%)
Risque cytogénétique à l'inclusion – Nombre de patients (%) ‡					
Risque élevé	100 (13%)	210 (23%)	12 (8%)	6 (6%)	5 (10%)

Caractéristiques	ASPIRE	ENDEAVOR	Registre d'hémopathies malignes régional	Registre d'hémopathies malignes régional	Registre d'hémopathies malignes régional
Risque standard	317 (40%)	575 (62%)	84 (57%)	57 (61%)	30 (61%)
Inconnu	375 (47%)	144 (15%)	52 (35%)	30 (32%)	14 (29%)
Clairance de la créatinine §					
Moyenne – ml/min	85,4±29,6	75,9±32,1	75,7±27,7	74,7±28,9	71,9±29,8
Distribution – Nombre de patients (%)					
<30 mL/min		56 (6%)	9 (6%)	8 (9%)	6 (12%)
30 à 50 ml/min	56 (7%)	128 (14%)	13 (9%)	8 (9%)	2 (4%)
≥ 50 ml/min	728 (92%)	745 (80%)	120 (82%)	74 (80%)	39 (80%)
Inconnu ou autre valeur	8 (1%)		6 (4%)	3 (3%)	2 (4%)
Sérum β2-microglobuline – Nombre de patients (%)					
< 2,5 mg/litre	77 (19%)		30 (20%)	21 (23%)	9 (18%)
≥ 2,5 mg/litre	638 (81%)		92 (62%)	55 (59%)	31 (63%)
Non disponible	-		26 (18%)	17 (18%)	9 (18%)

*Valeur haute – valeur basse. Valeurs sont des moyennes ± Ecart type

†Les scores de l'Eastern Cooperative Oncology Group (ECOG) vont de 0 à 5, avec 0 pour aucun symptôme.

‡ Le groupe à **risque élevé** est constitué de patients avec le sous-type génétique t (4; 14) ou t (14; 16) ou avec la délétion 17p dans 60% ou plus de cellules plasmatisques, selon le contrôle des échantillons de moelle osseuse obtenus au début de l'étude. Le groupe à **risque standard** est constitué de patients sans t (4; 14) ou t (14; 16) et avec délétion 17p dans moins de 60% des cellules plasmatisques. La valeur limite de 60% pour la proportion de cellules plasmatisques avec délétion 17p a été utilisé sur la base des recommandations du Myeloma Workshop international Consensus Panel 2.17

§ D'après les critères d'admissibilité, les patients devaient avoir une clairance de la créatinine d'au moins 50 ml par minute lors du screening. Un patient du groupe témoin a une clairance de la créatinine inférieure à 30 ml par minute au début de l'étude.

Un patient (0,3%) dans chaque groupe a reçu au moins 4 traitements.

ND Non Disponible

► La structure du modèle

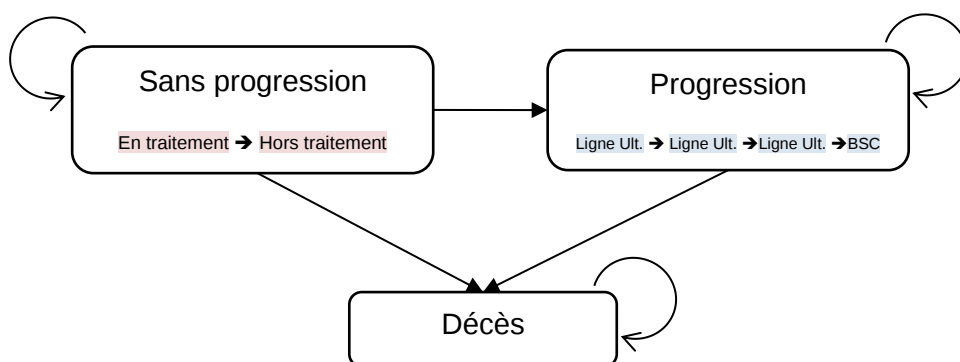
Type de modèle

Le modèle est un modèle de cohorte de type « aire sous la courbe » (*partitioned survival model*) à 3 états.

Ce choix repose sur une revue systématique des évaluations médico-économiques et rapports des agences d'évaluations (HTA) dans le myélome multiple, réalisée en février 2015. L'approche semi-markovienne est l'approche la plus utilisée dans le myélome multiple en rechute.

États du modèle

Trois états distincts et mutuellement exclusifs sont définis : « Sans progression » ; « avancé » et « décédé ».

Figure 1 : flux de patient dans le modèle

Remarque : Ligne Ult : Ligne ultérieure fondée sur 50% de PomD : IMNOVID® + faible dose de dexaméthasone; BendaD et 50% LEVACT® + dexaméthasone; BSC : Meilleurs soins de soutien

Événements intercurrents

Les arrêts de traitement dans l'état sans progression

Kyprolis® est administré jusqu'à un maximum de 18 cycles (règle d'arrêt dans l'essai ASPIRE), bien que le RCP mentionne une prescription possible au-delà.

L'impact d'un arrêt de traitement pour chacune des molécules intégrées dans les trois associations étudiées (KRd, Rd et Vd) est pris en compte :

- sur les coûts, en appliquant les rapports de risques instantanés observés dans les essais cliniques aux courbes de survie sans progression ;
- sur l'efficacité, par l'utilisation de données de survie sans progression incluant les patients ayant arrêté leur traitement avant progression⁵ (données en ITT).

Aucune modification des utilités n'a été considérée à l'arrêt du traitement avant la progression.

Les passages d'une ligne de traitement à une autre dans l'état post-progression

Plusieurs séquences de traitement sont considérées en post-progression, en fonction de la sous-population modélisée.

Tableau 3. Comparateurs et séquences de traitement par ligne utilisés dans le modèle mis à jour

Sous-population	2	3	4	5 ⁺	
2 ^{ème} ligne	KRd Rd Vd	Pom-d /benda-d	Benda-d /Pom-d	Traitement non spécifié	Soins palliatifs
3 ^{ème} ligne		KRd Rd Vd	Benda-d /Pom-d	Traitement non spécifié	Soins palliatifs
4 ^{ème} ligne			KRd Rd Vd	Traitement non spécifié	Soins palliatifs

Les effets indésirables

Avant progression, le modèle inclut des effets indésirables liés au traitement. Les événements retenus sont les EI de grade ≥ 3 observés dans l'essai ASPIRE, avec une fréquence d'occurrence

⁵ Seuls les patients ayant initié un autre traitement ont été censurés pour l'estimation de la survie sans progression.

≥ 5% pour une des stratégies thérapeutiques considérées ou avec une différence de 1 % entre les bras Rd et KRd.

Les diarrhées et les neuropathies périphériques, qui ont une fréquence importante dans le bras Vd, ont été ajoutées à la liste des événements sélectionnés.

Aucun effet indésirable n'est modélisé après progression de la maladie.

Principales hypothèses simplificatrices de modélisation

- Il est supposé que la survie globale tient compte de la mortalité toutes causes en raison du caractère avancé du myélome multiple.
- Les arrêts de traitement sont capturés par l'utilisation des données en ITT, et donc considérés uniquement au niveau des coûts.
- Le cross over n'a pas été considéré dans l'analyse (9 patients du bras KRd et 8 patients du bras Rd ont reçu Kyprolis® après la progression).
- Les traitements après progression n'ont aucun effet sur la survie globale et aucun effet indésirable n'est modélisé après progression.

Gestion de la dimension temporelle

- [1] Durée de simulation

Une durée maximale de simulation est fixée à 15 ans (durée de simulation de 195 cycles).

- [2] Durée des cycles dans le modèle

Un cycle de 4 semaines a été utilisé, correspondant à la durée des cycles dans le protocole de l'étude ASPIRE. Aucune correction de demi-cycle n'a été appliquée.

- [3] Méthodes d'extrapolation à long terme des données disponibles

❖ *Extrapolation de la survie globale (SG) et de la survie sans progression (SSP) associées au traitement Rd*

Une extrapolation des données de survie issues du Registre des Hémopathies Malignes a été réalisée sur 15 ans, sous l'hypothèse que ces données de survie représentent le traitement Rd et sont correctement ajustées par la fonction paramétrique de Weibull., sélectionnée selon la méthode ci-dessous.

L'ajustement d'une courbe paramétrique aux données a été réalisé sur une sélection de 20 points extraits de la courbe de Kaplan Meier pour la SG et la SSP par ligne de traitement répartis sur les 5 ans de suivi.

Les fonctions testées sont la distribution de Weibull, la distribution Log-logistique et la distribution exponentielle. Pour sélectionner la distribution permettant d'ajuster au mieux l'extrapolation aux données disponibles, l'industriel s'est fondé sur les critères d'information statistique d'Akaike (AIC) et de Bayes (BIC), sur une inspection visuelle de l'ajustement par rapport aux courbes de Kaplan-Meier issues du registre et sur une comparaison avec des données en vie réelle.

Pour la SG et la SSP, les critères statistiques sont en faveur de la distribution de Weibull pour la seconde ligne. La distribution log-logistique était la meilleure distribution d'un point de vue statistique, en 3^{ième} ligne et 4^{ième} ligne, avec cependant des survies modélisées bien supérieures à celles obtenues avec la Weibull. Afin d'être conservateur et dans une logique d'uniformité de la fonction entre les lignes, l'industriel a retenu la fonction de Weibull pour l'analyse de référence. Une analyse de sensibilité avec les autres fonctions a été réalisée.

Tableau 4 : Prédictions de la SG à 5, 8, 10 et 25 pour toutes les distributions paramétriques

	5 ans	8 ans	10 ans	25 ans
Projections				

2^{ème} ligne				
Weibull	31.6%	15.1%	9.2%	0.2%
Loglogistique	33.7%	21.4%	16.8%	5.6%
Exponentielle	32.6%	16.6%	10.6%	0.4%
3^{ème} ligne				
Weibull	19.3%	8.2%	4.8%	0.1%
Loglogistique	22.0%	13.7%	10.8%	3.7%
Exponentielle	16.4%	5.6%	2.7%	0.0%
4^{ème} ligne				
Weibull	12.9%	5.0%	2.7%	0.0%
Loglogistique	15.6%	9.5%	7.4%	2.6%
Exponentielle	7.4%	1.6%	0.6%	0.0%
Données en vie réelles				
Barlogie 2010	27% - 35%	15% - 20%	N/A	N/A
Pulte 2014	44.6%	NR	25%	N/A

❖ *Extrapolation de l'effet traitement de KRd vs Rd*

L'hypothèse est faite que le Hazard Ratio (HR) obtenu dans l'essai clinique ASPIRE capture correctement l'effet incrémental de KRd par rapport à Rd sur la survie globale et sans progression jusqu'à 36 mois.

Au-delà de cette période, en l'absence de preuves, les rapports de risque sont supposés égaux à 1.

❖ *Extrapolation de l'effet traitement de Vd vs Rd*

L'hypothèse est faite que le Hazard Ratio (HR) obtenu dans l'analyse statistique groupée des essais ASPIRE et ENDEAVOR capture correctement l'effet incrémental de Vd par rapport à Rd sur la survie globale et sans progression jusqu'à 18 mois.

Au-delà de cette période, en l'absence de preuves, les rapports de risque sont supposés égaux à 1.

► **L'estimation de la proportion de patients dans les états du modèle**

[1] Sources de données

Trois sources de données différentes sont mobilisées pour estimer les courbes de survie associées au bras RD (Registre des hémopathies malignes), le rapport des risques KRd vs Rd (essai ASPIRE) et le rapport des risques Vd vs Rd (analyse statistique groupée des essais ASPIRE et ENDEAVOR).

❖ *Le Registre des Hémopathies Malignes (Côte d'Or, de Gironde et du Calvados) est utilisé pour modéliser la survie du traitement de référence (Rd) par ligne de traitement.*

L'industriel privilégie le recours à des données de registre, à la place des données du bras Rd de l'essai ASPIRE, en raison du non réalisme, au regard de l'évolution naturelle de la maladie, des tendances observées dans l'essai ASPIRE en termes de survie globale (SG) et de survie sans progression (SSP) lorsque l'on passe d'une ligne de traitement à une autre⁶. L'industriel explique

⁶ Contrairement à ce qui est observé sur des données en vie réelles, il n'y a pas de tendance à une diminution de la SSP médiane de la stratégie Rd avec l'augmentation de la ligne de traitement (17.6 mois [15.0 ; 22.2] en ligne 2, 18.5 mois

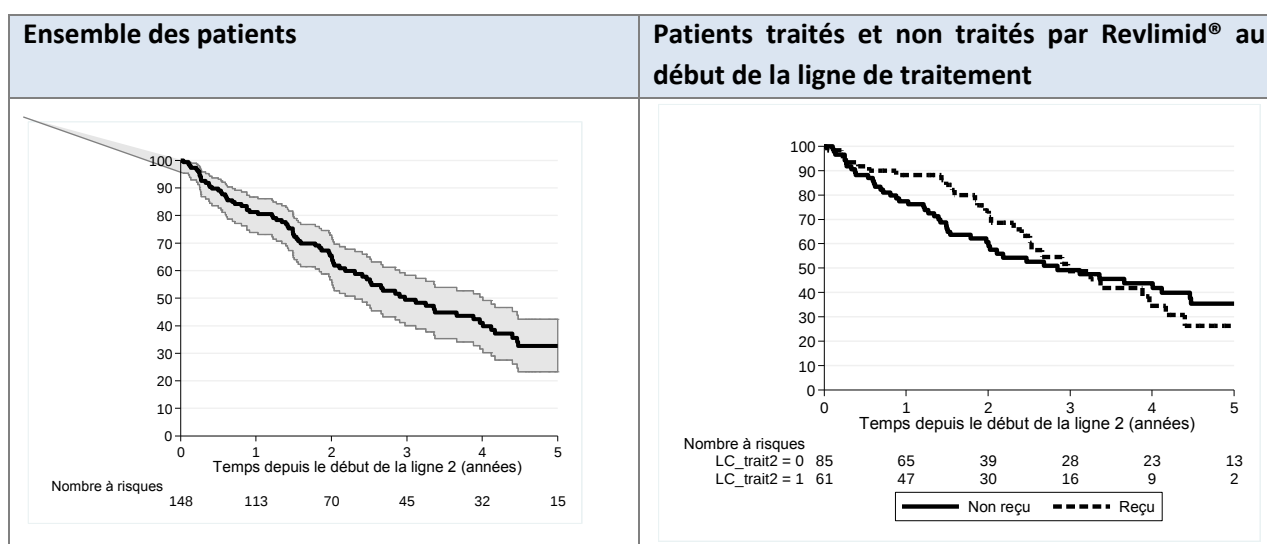
ces observations par l'inclusion dans l'essai de patients en 3^{ème} ou 4^{ème} ligne qui ont un meilleur pronostic que les patients en vie réelle.

L'industriel fait donc l'hypothèse que les données de survie observées dans le registre sont représentatives des survies attendues en vie réelle sous Rd, bien qu'en pratique seuls 44%, 49% et 33% des patients du registre ont reçu un traitement incluant Rd respectivement en lignes 2, 3 et 4.

La Figure 2 montre que les survies globales observées pour le groupe de patients Rd sont incluses dans l'intervalle de confiance de la survie pour l'ensemble des patients. Par ailleurs, l'utilisation des données totales permet d'augmenter la précision de l'extrapolation, la taille d'échantillon pour les patients Rd uniquement étant particulièrement restreinte pour les lignes 3 et 4. Enfin, l'utilisation des données de registre permettrait de garantir la représentativité de la population simulée par rapport à la population cible française.

Une analyse de sensibilité utilisant la borne supérieure de l'estimation de la SG et de la SSP, qui s'approche des données observées dans le bras Rd, est présentée afin d'évaluer l'impact d'utiliser les données globale à la place des données du bras Rd.

Figure 2 - Données de SG du registre des hémopathies malignes pour les patients en 2^{ème} ligne



❖ *L'essai ASPIRE est utilisé pour estimer les rapports de risques SSP et SG entre les bras de traitement KRd et Rd.*

ASPIRE est un essai de phase III comparant KRd à Rd, chez 792 patients adultes atteints d'un myélome multiple en rechute après 1 à 3 traitements. Le temps de suivi est de 32 mois.

Une réduction du risque de progression ou de décès de 31 % (HR=0,690 ; IC95 % [0,570 ; 0,834], p=0,0001) a été observée en faveur de KRd, correspondant à une amélioration de la médiane de SSP de +8,7 mois en valeur absolue (26,3 mois versus 17,6 mois).

Un hazard ratio de survie globale en faveur du bras KRd (HR= 0,79 ; IC95 % [0,63 ; 0,99], p=0,04) et un nombre moins élevé de décès dans le bras KRd versus bras Rd (143 [36,1 %] versus 162 [40,9 %]).

A la date de cut-off, les médianes de SG ne sont atteintes dans aucun des deux bras. L'étude se poursuit pour disposer de données matures de survie globale.

[14,0 ; 25,0] en ligne 3 et 14,8 mois [10,0 ; 22,1] en ligne 4) ou à une augmentation du nombre de décès (42,0% en ligne 2, 38% en ligne 3 et 28,8% en ligne 4).

Le cross-over n'a pas été considéré dans l'analyse car il est très équilibré entre les deux bras.

- ❖ *L'analyse groupée des essais ASPIRE et ENDEAVOR est utilisée pour estimer les rapports de risques SSP et SG entre les bras de traitement Vd et Rd.*

ENDEAVOR est un essai de phase III comparant Kd à Vd, chez 929 patients adultes présentant un myélome multiple en rechute après au moins une ligne de traitement antérieure. Le temps de suivi médian est de 12 mois.

Une réduction statistiquement significative du risque de progression ou de décès de 47 % (HR=0,533 ; IC95 % [0,437 ; 0,651], $p < 0,0001$) dans le bras Kd par rapport au bras Vd a été observée. Le p est inférieur à la borne spécifiée au protocole pour un arrêt prématuré de l'étude lors de l'analyse intermédiaire de la SSP (pour rappel p spécifié = 0,0096). Selon les courbes de Kaplan-Meier, la médiane de SSP est de 18,7 mois (IC95 % [15,6 ; non estimable]) dans le bras Kd versus 9,4 mois (IC95 % [8,4 ; 10,4]) dans le bras Vd, soit un doublement de la SSP avec une amélioration de +9,3 mois en valeur absolue en faveur du bras Kd.

[2] Méthode d'estimation de la proportion de patients par état du modèle

- La proportion de patients dans l'état « survie sans progression » (% SSP) est estimée directement par la méthode de l'aire sous la courbe.
- La proportion de patients dans l'état « survie post progression » (% SPP) est estimée indirectement par la différence entre la courbe de survie globale et la courbe de survie sans progression.
 $\% \text{ SPP} = \% \text{ SG} - \% \text{ SSP}$.
- La proportion de patients dans l'état « décès » est estimée à partir du calcul de l'aire sous la courbe de survie globale.
 $\% \text{ Décès} = 1 - \% \text{ SG}$.

[3] Estimation des courbes de survie sans progression et de survie globale

- ❖ *Estimation des courbes de survie sans progression et de survie globale pour le traitement KRd*

Les courbes SSP et SG associées au traitement KRd sont dérivées de l'application d'un Hazard Ratio ($\text{HR}_{\text{KRd vs Rd}}$), capturant l'effet incrémental de KRd par rapport à Rd, aux courbes de survies retenue dans le bras Rd.

Les HR ont été estimés à partir de l'essai clinique ASPIRE, par des analyses de Cox multivariées, permettant d'ajuster l'effet du traitement sur les caractéristiques de base des patients, susceptibles d'avoir un impact sur la SG ou la SSP⁷. Ce modèle repose sur l'hypothèse des risques proportionnels de l'effet du traitement au cours du temps, lequel est testé par les auteurs.

Les effets du traitement étaient de HR=0,641 pour la SSP et HR=0,742 pour la SG, et ils ont été appliqués dans l'analyse de référence (Tableau 5) pour toutes les lignes de traitement.

L'effet traitement est donc supposé indépendant de la ligne de traitement. Selon l'industriel, les HR non-ajustés confirment que l'application d'une efficacité constante selon les lignes de traitement est appropriée

- pour la SG [$\text{HR}_{\text{ligne1}}=0,815$ (0,582 ; 1,143) ; $\text{HR}_{\text{ligne2}} = 0,788$ (0,523 ; 1,188) ; $\text{HR}_{\text{ligne3}} = 0,731$ (0,0466 ; 1,147)].
- pour la SSP [$\text{HR}_{\text{ligne1}}=0,713$ (0,532 ; 0,957) ; $\text{HR}_{\text{ligne2}} = 0,745$ (0,536 ; 1,036) ; $\text{HR}_{\text{ligne3}} = 0,682$ (0,468 ; 0,995)].

⁷ Les covariables sélectionnées, via une méthode de sélection automatique stepwise (avec seuil de 20% pour inclure une covariable et un seuil de 10% pour l'exclure), reportées dans le rapport technique.

Tableau 5. Rapports de risques instantanés entre KRd et Rd pour SSP et OS, appliqués sur les 36 premiers mois.

	Analyse multivariée				Analyse non ajustée			
	HR KRd vs Rd			p-value (bilatérale)	HR KRd vs Rd			p-value (bilatérale)
	Estimation	IC à 95 %			Estimation	IC à 95 %		
SSP	0,641	0,526	0,781	<0,0001	0,69	0,57	0,83	<0,0001
OS	0,742	0,587	0,938	0,0128	0,79	0,63	0,99	0,0018

Abréviations: Ic = intervalle de confiance; Kyprolis® KRd = () + lénalidomide (Lenalidomide®) + faible dose de dexaméthasone; Rd = lénalidomide (Kyprolis®) + faible dose de dexaméthasone.

❖ *Estimation des courbes de survie sans progression et de survie globale pour le traitement Vd*

Les courbes SSP et SG associées au traitement Vd sont dérivées de l'application d'un Hazard Ratio ($HR_{Vd \text{ vs } Rd}$), capturant l'effet incrémental de Vd par rapport à Rd, aux courbes de survies retenue dans le bras Rd.

Les HR sont estimés à partir d'une analyse groupée des données individuelles issues des essais ASPIRE (stratégies KRd et Rd) et ENDEAVOR (stratégie Vd), qui permet d'ajuster les résultats sur les caractéristiques individuelles des patients. La méthode est détaillée en annexe du rapport technique de l'industriel. L'industriel recourt à cette méthode car Krd, Rd et Vd ne peuvent être connectés dans un réseau en méta-analyse, sans utiliser des hypothèses cliniques fortes, notamment sur l'effet de la dexaméthasone.

Deux différences entre les essais ASPIRE et ENDEAVOR sont rapportées :

- Les populations sont similaires à l'exception des traitements antérieurs⁸ et de la fonction rénale, reflet des différences de critères d'inclusion des études ;
- Le suivi moyen est de 32 mois pour ASPIRE et 12 mois pour ENDEAVOR.

Les effets du traitement étaient de $HR=0,521$ pour la SSP et $HR=0,884$ pour l'OS, et ils ont été appliqués dans l'analyse de référence (Tableau 6) pour toutes les lignes de traitement.

L'hypothèse d'un HR constant dans le temps a été testée en utilisant un modèle de Cox avec l'interaction temps*traitement. Les résultats des tests d'interaction indiquent que les risques ne sont probablement pas proportionnels pour la SSP, en raison de la durée de suivi limitée dans l'étude ENDEAVOR. Cependant, l'industriel n'a pas retenu cette analyse, permettant d'estimer des HR différents selon la période, en raison du nombre de censures particulièrement important dans les derniers mois d'ENDEAVOR. L'estimation de la fonction de risque sur les différentes périodes est associée à un niveau d'incertitude important (>50% pour la SG et >35% pour la SSP au moment de l'analyse principale). L'impact de ce choix est testé en analyse de sensibilité.

Tableau 6. Rapports de risques instantanés entre Rd et Vd pour SSP et OS, appliqués sur les 18 premiers mois.

	Analyses multivariée			Analyse multivariée par morceau	
	HR Rd vs Vd			HR Rd vs Vd	
	Estimation	IC à 95 %		Estimation 0-12 mois	Estimation 12+ mois
SSP	0,521	0,417	0,651	0,5759	0,3423
OS	0,884	0,639	1,221	1,0261	0,6103

⁸ Les auteurs ont réalisé un test global (en rapportant les valeurs de p) de la différence des distributions pour chaque caractéristique des patients et il en ressort une différence sur l'exposition précédente au lenalidomide et au bortezomib entre les 2 études.

► L'estimation des probabilités d'occurrence d'un événement intercurrent

Estimation des probabilités d'arrêt de traitement avant progression

La durée sous traitement a été modélisée séparément pour chaque composante des associations, en appliquant les rapports de risques instantanés observés dans les essais ASPIRE et ENDEAVOR aux courbes de survie sans progression.

Les HR ont été calculés à partir de modèles de régression de Cox comparant la SSP avec le temps jusqu'à l'arrêt. Les données sont reportées ci-dessous.

Données introduites dans le modèle

	Rapport de risque
KRd (K)	HR=0,762 ; IC95% [0,596-0,973].
KRd ®	HR=0,723 ; IC95% [0,603-0,866].
KRd (d)	HR=0,689 ; IC95% [0,575-0,825].
Rd ®	HR=0,747 ; IC95% [0,628-0,888].
Rd (d)	HR=0,706 ; IC95% [0,594-0,839].
Vd (V)	HR=0,639 ; IC95% [0,543-0,753].
Vd (d)	HR=0,623 ; IC95% [0,528-0,733].

Estimation des probabilités d'effets indésirables

La probabilité d'être victime d'un EI par cycle (Tableau 7) est calculée à partir de la fréquence d'occurrence des EIs pendant toute la période de l'essai et de la durée moyenne du traitement à l'aide de la formule suivante :

Probabilité d'EI par cycle = $1 - [(1 - \text{Fréquence d'EI au cours de l'essai})^{(1/\text{durée moyenne du traitement})}]$.

Selon l'industriel, cette méthode est classique dans les modélisations, mais elle peut conduire à une sous-estimation de l'ensemble des EIs, en raison de l'absence de précision sur l'itération de certains événements indésirables⁹.

Deux hypothèses ont permis d'introduire les données sur les événements indésirables dans le modèle :

- les EI peuvent se produire avec un risque constant et sont distribués uniformément sur toute la durée du traitement ;
- les probabilités d'occurrence sont indépendantes du niveau de réponse au traitement, et de la ligne de traitement.

Une analyse de sensibilité a été réalisée en appliquant les taux d'EIs observés dans les essais lors du premier cycle avec la fréquence rapportée dans les essais cliniques.

Tableau 7. Fréquences de grade 3+ événements indésirables par cycle

Événement indésirable	KRd	Rd	Vd
Blood and Lymphatic System Disorders			
Neutropenia	1.42%	1.50%	0.18%
Anaemia	0.40%	0.55%	0.60%
Thrombocytopenia	0.70%	0.60%	1.12%
Leukopenia	0.12%	0.21%	0.06%
Febrile Neutropenia	0.09%	0.04%	0.06%
Lymphopenia	0.12%	0.09%	0.24%
Gastrointestinal Disorders			

⁹ Dans les essais, un patient n'est compté qu'une seule fois pour un même EI.

Diarrhoea	0.07%	0.12%	0.84%
Nausea	0.01%	0.04%	0.09%
Constipation	0.01%	0.02%	0.20%
Vomiting	0.00%	0.00%	0.03%
Vascular Disorders			
Thrombo-Embolic Event	0.08%	0.06%	0.03%
Respiratory, Thoracic and Mediastinal Disorders			
Dyspnoea	0.07%	0.02%	0.20%
General Disorders and Administration Site Conditions			
Fatigue*	0.27%	0.27%	0.78%
Metabolism and Nutrition Disorders			
Hypophosphataemia	0.24%	0.07%	0.09%
Hypokalaemia	0.16%	0.10%	0.09%
Nervous System Disorders			
Neuropathy Peripheral	0.07%	0.09%	0.72%
Infections and Infestations			
Pulmonary infection **	0.27%	0.22%	0.27%

Source: Calculé à partir des tableaux 14.3.3.2 (CSR ASPIRE) et 14.3.3.2 (CSR ENDEAVOR)

► Validation du modèle

Le modèle a été validé afin d'assurer la validité apparente (plausibilité) et la validité technique (vérification).

Les auteurs ont également comparé les données simulées et observées dans le modèle à partir de l'étude ASPIRE et du registre. Par ailleurs, certains résultats obtenus avec le modèle ont été comparés avec d'autres données publiées ou issus de la *chart review* menée par Amgen.

4.4.2 Analyse critique concernant la modélisation

► La population simulée

D'une manière générale, les patients du Registre des Hémopathies Malignes, bien que reposant sur des données départementales et non nationales (Côte d'Or, Gironde et Calvados), sont plus représentatifs de la population française que ceux des essais cliniques de phase III. En effet, les essais ont inclus des patients ayant un meilleur pronostic que les patients attendus en pratique courante.

► La structure du modèle

Le choix du modèle repose sur une revue de littérature sur les modèles soumis aux agences d'HTA qui est bien menée.

Le modèle choisi est conforme aux modèles classiquement retenus en oncologie.

Les hypothèses présentées par les auteurs n'appellent pas de commentaire.

► Gestion de la dimension temporelle

La durée des cycles du modèle est cohérente avec les séquences de traitements et l'histoire naturelle de la maladie.

Le choix d'une durée de simulation maximale de 15 ans n'est pas étayé par des données sur l'espérance de vie en pratique courante. Par ailleurs, le choix d'une durée fixe, quelle que soit la ligne de traitement, aurait pu être justifiée, considérant la réduction de la survie globale observée à chaque ligne de traitement supérieure.

Les méthodes utilisées pour extrapoler les données posent plusieurs questions.

- Concernant l'extrapolation des données du registre dans le bras Rd, le choix des fonctions paramétriques testées n'est pas justifié et reste assez restreint (seules trois fonctions testées). L'intérêt de retenir la même fonction paramétrique n'est pas argumenté, alors qu'il conduit à retenir la loi de Weibull qui ne correspond aux critères statistiques que sur 1 ligne de traitement. Cependant, les analyses de sensibilité montrent que le choix de la fonction de Weibull est défavorable par rapport aux deux autres fonctions testées.
- L'extrapolation jusqu'à 36 mois de l'effet incrémental de KRd par rapport à Rd repose sur le protocole de l'essai ASPIRE.
- L'extrapolation jusqu'à 36 mois de l'effet incrémental de KRd par rapport à Rd et jusqu'à 18 mois de l'effet incrémental de Vd par rapport à Rd peut générer un biais. L'impact de ce choix n'est pas évalué en analyse de sensibilité, contrairement à ce qui est énoncé par les auteurs. Il n'est pas possible de savoir si Vd serait toujours dominé avec une durée de persistance de l'effet traitement identique pour KRd et Vd.

► Estimation de la proportion de patients dans les états du modèle

❖ Estimation pour le traitement de référence (Rd)

La principale source de discussion porte sur le choix de la source de données retenue pour modéliser les courbes de survie associées au traitement de référence (Rd). Dans le registre, moins de 50% des patients reçoivent Rd, quelle que soit la ligne de traitement. Les courbes de survies ne représentent donc pas Rd mais le traitement au choix du médecin (TCM) en cours en France sur la période considérée.

Cependant, l'analyse des données disponibles dans la source alternative (bras Rd de l'essai ASPIRE) indique que les courbes de survie par ligne qui en seraient issues ne seraient pas réalistes en raison d'une population incluse avec un meilleur pronostic que celui attendu en pratique courante. Le choix de modéliser les courbes de survie du traitement de référence à partir du registre et non de l'essai ASPIRE est donc acceptable dans ce cas particulier. L'analyse de sensibilité montre que le choix du registre conduit à un RDCR plus favorable au produit que le choix de l'essai ASPIRE dans les trois lignes de traitement (RDCR $\nearrow$ +11% à +27%).

Il reste que les courbes de survie ont été modélisées sur l'ensemble de la population du registre et non sur la population recevant Rd. Bien que les arguments avancés par l'industriel soient recevables, les analyses de sensibilité montrent que le choix de la population globale du registre conduit à un RDCR plus favorable au produit que le choix de la population recevant Rd dans les trois lignes de traitement (RDCR $\nearrow$ +9% à +1%).

Le choix de sources différentes pour estimer le risque de base sur Rd et l'effet traitement relatif de KRd et Vd pose un autre problème, puisque cela conduit à appliquer des HR estimés sur une population à une courbe de survie observée dans une autre population. Ceci peut constituer un biais si les risques de bases (variables pronostiques) et les caractéristiques (variables prédictives, facteurs de confusion, biais écologiques) sont trop différents entre les populations (registre vs. essais). Or, les auteurs ne font pas la démonstration d'une homogénéité des populations entre le registre et les essais cliniques en termes de distribution des variables pronostiques et prédictives.

En conclusion, la modélisation des fonctions de survie utilisées dans le bras de référence est une source d'incertitude.

❖ Estimation de l'effet traitement de Krd et Vd

Le modèle repose sur l'hypothèse d'un effet traitement équivalent entre les lignes pour KRd et tous les autres comparateurs du modèle.

Pour seule justification, les auteurs présentent les $HR_{KRd \text{ vs } Rd}$ non-ajustés, pour SSP et SG par ligne de traitement, et les jugent équivalents sans pour autant tester cette équivalence. Des analyses de sensibilité auraient pu être réalisées pour quantifier l'impact sur les RDCR d'une variation des $HR_{KRd \text{ vs } Rd}$ entre 0,815 et 0,713 pour la SG, et entre 0,713 et 0,682 pour la SSP.

❖ Estimation de l'effet traitement de Krd

Les HR de Krd vs Rd sont issus d'un modèle de Cox multivarié, c.-à-d. des résultats de l'essai de phase III ASPIRE ajustés sur des covariables individuelles. Le recours à une analyse multivariée au lieu des données directement issues de l'essai n'est pas suffisamment justifié et l'impact de ce choix n'a pas fait l'objet d'une analyse de sensibilité permettant de tester l'impact des HR issus directement de l'essai clinique.

Une méthode de sélection automatique des co-variables est réalisée. La méthode est détaillée en annexe du rapport technique incluant les variables étudiées ainsi que les p-values. Pour certaines variables, les p-values diffèrent selon des sous-groupes qui ne sont pas définis.

❖ Estimation de l'effet traitement de Vd

L'estimation de l'effet relatif de Vd repose sur une méthode dont le niveau de preuve est inférieur à celui d'une méta-analyse (p.ex. perte de randomisation ; ajustement uniquement sur les facteurs connus, mesurés et disponibles dans les essais ; nombreux biais notamment écologiques). Ce choix est cependant acceptable, compte tenu de la fragilité du réseau et des hypothèses nécessaires à l'inclusion de Vd dans le réseau.

Pour autant, l'estimation de l'effet traitement de Vd pose plusieurs questions, susceptibles de remettre en cause la robustesse du HR introduit dans l'analyse de référence.

Le choix de la méthode statistique utilisée par les auteurs est insuffisamment argumenté. En effet, les auteurs du rapport signalent différentes méthodes de comparaison indirecte avec ajustement, telles que les STC (*Simulated treatment comparison*) ou MAIC (*Matching-adjusted indirect comparison*). Ils ne présentent pas de références pour justifier et expliquer la méthodologie finalement retenue.

En théorie, ces méthodes peuvent utiliser les données individuelles d'un ou plusieurs essais, ainsi que des données agrégées. Les efficacités relatives sont alors mesurées en ajustant sur les facteurs de confusion potentiels et la comparabilité des groupes est rétablie avec des scores de propension. L'industriel n'a pas inclus de données agrégées en utilisant des scores de propension afin d'inclure d'autres comparateurs, et ce sans justification.

Bien que les auteurs concluent que les essais ASPIRE et ENDEAVOR sont globalement comparables¹⁰ et que les différences dans les caractéristiques de base seront correctement contrôlées dans leurs analyses, il est important de noter trois limites fortes dans ce dossier.

- Les temps de suivi entre les deux essais sont très différents : 12 mois pour ENDEAVOR (bras Vd 11.9 mois ; bras Kd 12.5 mois) et 32 mois pour ASPIRE (bras Rd 31.5 mois ; bras KRd 32.3 mois). Ces différences entraînent des incertitudes sur la prédiction à long terme pour la survie globale, qui est une des données d'entrée les plus sensibles du modèle.
- Les données de SSP sont matures (critère analyse intermédiaire), mais les données sont immatures pour l'OS, surtout pour ENDEAVOR.
- Certaines variables individuelles n'ont pas pu être incorporées, soit parce qu'elles ne sont pas présentes dans les 2 essais, soit parce qu'elles ne sont pas sous le même format. Par exemple, et contrairement à ce qui est dit par les auteurs, l'analyse statistique menée ne permet pas d'ajuster sur les traitements reçus précédemment, qui sont pourtant un critère de stratification de la randomisation.

Un autre élément méthodologique est également critiquable : l'application d'un HR constant dans le modèle en analyse de référence, alors que l'hypothèse de proportionnalité des risques n'est pas vérifiée dans l'essai ENDEAVOR pour la SSP. En théorie, le modèle de Cox ne devrait donc pas

¹⁰ Ces 2 essais ont des designs similaires : essais de phase III, multicentriques, incluant des patients atteints de myélome multiple, en rechute, qui ont déjà eu entre 1 et 3 thérapies antérieures. Les critères d'inclusions et de non-inclusion, ainsi que les facteurs de stratification sont décrits.

s'appliquer dans ce cas. Néanmoins, au vu de la fragilité de l'évaluation du HR par morceaux sur les derniers mois de l'essai et de l'absence d'impact observé dans les analyses de sensibilité, le choix des auteurs semble être acceptable.

En conclusion, la méthode proposée pour inclure Vd est une analyse statistique ajustée de deux essais, avec ajustement sur données individuelles. Ce type d'analyse a un niveau de preuve équivalent à celui d'études observationnelles. **L'estimation du HR pour Vd n'est pas fiable**, notamment à cause d'un temps de suivi entre ASPIRE et ENDEAVOR différent, et des bénéfices de la randomisation non conservés. On ne connaît pas le sens du biais (surestimation ou sous-estimation), surtout que les HR semblent dépendants du temps.

► **L'estimation des probabilités d'occurrence d'un événement intercurrent**

Estimation des probabilités d'arrêt de traitement avant progression

L'évaluation médico-économique présentée évalue l'efficience de KRd avec les règles d'arrêt de traitement imposées dans l'essai, et non pas les possibilités offertes par le RCP. Il n'est pas exclu que le traitement soit prescrit sur des durées supérieures à 18 cycles en vie réelle, ce qui aurait pour conséquence d'augmenter les coûts de traitement et donc le RDCR. Il est à noter que le bénéfice clinique potentiel associé à une durée de traitement plus longue n'a pas été étudié.

La comparaison des durées de traitement estimées dans le modèle avec des données en vie réelle montre une surestimation des durées de traitement dans le modèle, surtout pour REVLIMID® (24 mois vs. 12 mois). Les analyses de sensibilité ont permis de mettre en évidence que ce paramètre influence les $RDCR_{KRd \text{ vs } Rd}$, quelle que soit la ligne de traitement : une surestimation de la durée de traitement pour RD dans les bras KRd et Rd est défavorable au produit Kyprolis®.

Estimation des probabilités d'effets indésirables

Dans l'étude ASPIRE, une toxicité plus importante a été rapportée dans le groupe KRd en termes de troubles cardiaques, un risque identifié comme important dans le plan de gestion de risques, lesquels ne sont pas repris par les auteurs dans la modélisation.

► **Validation du modèle**

Les méthodes de validation mise en œuvre sont décrites.

4.5 Mesure et valorisation des états de santé

4.5.1 Evaluation des résultats de santé telle que présentée par les auteurs

► Méthode et données

Utilité dans l'état sans progression

Une revue systématique de la littérature a été réalisée afin d'identifier les données d'utilité disponibles dans le myélome. Deux études observationnelles et un essai clinique ont été analysés et rejetés après analyse critique.

La méthode retenue est l'utilisation des données de qualité de vie collectées dans l'essai ASPIRE à partir du questionnaire QLQ-C30, après leur conversion en score d'utilité EQ-5D par un algorithme de mapping.

Le choix de l'algorithme de mapping repose sur une revue de la littérature, visant à identifier les algorithmes associés au questionnaire QLQ-C30. Deux publications ont été identifiées (Proskorovsky et al. 2014 ; Versteegh et al. 2012). L'algorithme de Proskorovsky a été retenu par l'industriel pour 3 raisons : son développement, qui a été fait sur la base des données de patients atteints de myélome multiple, la proximité de la population ayant servi à élaborer l'algorithme avec la population simulée dans le modèle et, enfin, son utilisation dans une soumission au NICE.

❖ *Valeurs d'utilité avant et après progression pour KRd et Rd*

Pour estimer la valeur d'utilité avant progression pour KRd et Rd, les utilités estimées par mapping ont été agrégées entre les cycles 1 et 18 pour chacun des bras de traitement. Les valeurs obtenues par moyenne pondérée étaient de 0.7390 pour le bras KRd, 0.7220 pour le bras Rd.

La désutilité liée à une progression a été estimée comme étant la moyenne des différences de l'utilité avant et après la progression. L'utilité après progression est estimée en mettant en commun les deux bras KRd et Rd (soit 0.7320). Cette méthode est justifiée par l'industriel par le nombre limité de patients ayant progressé dans l'essai ASPIRE et pour lesquels le QLQ-C30 était disponible (N=66 et 40 dans les bras KRd et Rd respectivement). La désutilité liée à la progression est estimée à 0.050.

Pour obtenir les valeurs d'utilité par ligne, l'industriel a appliqué les utilités obtenues pour ASPIRE à la sous-population en 3^{ème} ligne de traitement. La valeur de désutilité liée à la progression a été ajoutée à celle de la 3^{ème} ligne afin d'obtenir celle de la 2^{ème} ligne et soustraite à celle de la 3^{ème} ligne afin d'obtenir celle de la 4^{ème} ligne.

❖ *Valeurs d'utilité avant et après progression pour Vd (analyse de référence) et les autres comparateurs (analyse de scénario).*

Il est supposé que l'utilité avant progression et post-progression est la même dans tous les bras de traitement.

Selon l'industriel, l'hypothèse d'une utilité avant progression associée à Vd identique à l'utilité de Rd peut conduire à surestimer la valeur d'utilité associée à Vd, dans la mesure où le taux de réponse à Vd est inférieur à Rd et que son mode d'administration est plus contraignant.

L'hypothèse de valeurs d'utilité équivalentes entre les bras de traitement est supposée conservatrice par l'industriel quel que soit le comparateur (Vd, Td, Pom-d et Benda).

Tableau 8 : Données d'utilité dans l'analyse de référence (3^{ème} ligne)

Données introduites dans le modèle					
Pré-progression			Post-progression		
KRd	Rd	Vd	KRd	Rd	Vd
0,7390	0,7220	0,7220	0,6814	0,6814	0,6814

Désutilité liée au traitement

Dans l'analyse de référence, l'industriel a considéré que les utilités avant progression, calculées à partir de l'étude ASPIRE sur les données spécifiques à chacun des traitements KRd et Rd, prenaient en compte le niveau de réponse, les événements indésirables et les conséquences du mode d'administration du traitement sur la qualité de vie.

En conséquence, aucune désutilité supplémentaire n'a été introduite dans le modèle pour prendre en compte l'impact des EIs sur la qualité de vie. Une analyse de sensibilité permet de tester l'impact de cette hypothèse.

Le passage d'une ligne de traitement à une autre n'entraîne aucun impact sur l'utilité. Une analyse de sensibilité permet de tester l'impact de cette hypothèse.

Autres désutilités

Aucun impact du vieillissement des patients sur l'utilité n'a été considéré.

► **Résultats**

Quelle que soit la ligne de traitement, KRd est associé aux survies sans progression et survie globale les plus longues. Malgré des survies globales relativement similaires, Vd est associé à une SSP moyenne inférieure à Rd.

Les QALYs gagnés proviennent essentiellement du gain en années de vie. On observe toutefois un gain de QALY inférieur au gain d'année de vie et décroissant au fur et à mesure des lignes.

Tableau 9. Efficacité actualisée : Analyse de référence

Traitements	SSP	LY	QALY
2^{ème} ligne			
KRd	3,41	4,14	3,17
Vd	1,60	3,55	2,51
Rd	2,57	3,68	2,71
3^{ème} ligne			
KRd	2,16	3,18	2,25
Vd	0,91	2,51	1,68
Rd	1,55	2,66	1,83
4^{ème} ligne			
KRd	1,62	2,59	1,73
Vd	0,57	1,92	1,24
Rd	1,09	2,08	1,36

Validation des simulations produites

Selon l'industriel, il est probable que l'estimation de l'utilité à partir des données de l'essai ASPIRE surestime la qualité de vie des patients jusqu'à la fin de vie (p. 78 du rapport technique).

4.5.2 La comparaison des simulations à 2 ans de KRd et Rd aux données de survie observées dans l'essai ASPIRE pour les bras KRd et Rd est raisonnable pour la cohorte de patients en 2^{ème} ligne. Analyse critique de l'estimation des résultats de santé

► **Méthode et données**

La méthode de valorisation de la qualité de vie dans le modèle soulève plusieurs critiques.

Premièrement, le recours aux données issues de l'essai ASPIRE soulève deux critiques :

- d'une part, il est nécessaire de transformer les données recueillies par une méthode de mapping. Ces méthodes, bien qu'acceptées par la HAS en cas d'absence de données directement re-

cueillies par EQ-5D, génèrent une incertitude importante. Les analyses de sensibilité présentées selon un algorithme différent, qui dégrade les valeurs d'utilités avant progression, montre que le choix de l'industriel est favorable au produit pour toutes les sous-populations (RDCR $\geq +3\%$ à 7%).

- d'autre part, l'essai ASPIRE inclut des patients plus jeunes et avec un pronostic meilleur à celui attendu en vie réelle. Ceci peut avoir eu pour conséquence de surestimer la qualité de vie, en particulier après progression de la maladie et sur les lignes de traitement supérieures (impact inconnu).

Deuxièmement, l'hypothèse d'une valeur d'utilité en L3 correspondant à la valeur d'utilité estimée dans l'essai ASPIRE est insuffisamment justifiée si l'on considère que 43,1% des patients inclus le sont en 2^{ème} ligne.

Troisièmement, l'hypothèse, récurrente dans les modèles, d'un impact sur la qualité de vie des EI englobé dans les estimations spécifiques de l'utilité par traitement, conduit probablement à sous-estimer l'effet des EI dans la mesure où le recueil de données est disjoint du calendrier de survenue des EI (idem pour les hospitalisations). Cependant, l'analyse de sensibilité réalisée par les auteurs ne met pas en évidence d'impact sur le RDCR.

Quatrièmement, l'impact de la progression de la maladie sur la qualité de vie est insuffisamment pris en compte. D'une part, le passage d'une ligne de traitement à une autre entraîne un décrétement d'utilité constant. Cependant, ce choix semble défavorable au produit puisque l'analyse de sensibilité incluant une désutilité supplémentaire à partir de la 5^{ème} ligne de traitement conduit à une réduction du RDCR (RDCR $\searrow -8\%$ à -1%). D'autre part, aucun impact du vieillissement des patients sur l'utilité n'a été considéré. L'absence d'analyse de sensibilité ne permet pas de connaître l'impact de cette simplification sur le RDCR.

Cinquièmement, l'hypothèse d'une similarité des valeurs d'utilités pour tous les comparateurs (valeurs identiques à Rd) génère de l'incertitude sur l'évaluation des QALYs sans qu'elle ne soit explorée en analyse de sensibilité.

En conclusion, il est vraisemblable que la méthodologie utilisée sous-estime l'impact de la pathologie sur la qualité de vie. Les analyses de sensibilité ont permis de mettre en évidence que cela surestime l'efficience de Kyprolis®.

Les analyses ont également mis en évidence que l'impact de la valorisation des utilités est complexe, et qu'une mésestimation de la qualité de vie ne joue pas de la même façon selon la sous-population et selon qu'elle porte sur la pré ou la post progression. Cela est vraisemblablement dû aux différences de durées dans les deux phases de la maladie, selon les sous-populations et selon les bras de traitement. Malheureusement, l'absence d'interprétation par l'industriel des nombreuses analyses de sensibilité réalisées sur les valeurs d'utilité n'a pas permis d'aller plus loin dans les conclusions.

► Résultats

Le modèle surestime possiblement la survie globale et la survie sans progression, dans une mesure relative à moyen terme, mais de manière plus forte et problématique à long terme.

En plus du choix de la méthode d'ajustement utilisée pour extrapoler les données, le modèle semble mal rendre compte de la dégradation de l'état de santé des patients après la seconde ligne.

4.6 Mesure et valorisation des coûts

4.6.1 Evaluation des coûts telle que présentée par les auteurs

Les coûts pris en compte dans le modèle sont :

- les coûts associés aux traitements (coût d'acquisition, coût d'administration, coût de prise en charge des événements indésirables) ;

- les coûts de prise en charge associés aux hospitalisations (suivi de la maladie et soins palliatifs) et au suivi ambulatoire.

► **Mesure, valorisation et calcul des coûts**

Source de données

Une revue de la littérature n'a pas permis d'identifier une source fiable pour documenter les ressources consommées dans le myélome multiple.

L'évaluation des ressources non liées au traitement repose sur l'étude « healthcare resource use » qui est une analyse rétrospective de 502 dossiers patients en France (AMGEN, 2015). Cette étude analysait les dossiers médicaux de patients en rechute (de la deuxième à la cinquième ligne) suivis par différents médecins (dont 61% d'hématologues) dans 76 hôpitaux différents à travers la France.

Seules les données observées au cours de la phase d'administration des traitements ont été utilisées, en raison de durées d'observation plus longues et de taille d'échantillon plus importante.

Les consommations de soins ont été valorisées grâce aux derniers référentiels de remboursement en vigueur: Etude Nationale Commune de Coûts (ENCC) 2013, Nomenclature Générale des Actes Professionnels (NGAP) 2013, Nomenclature des actes de biologie médicale (NABM) 2013, base nationale des médicaments. Les coûts de transport ont été extraits d'une étude publiée en 2014.

Tableau 10. Volumes consommés et coûts unitaires

Ressources	Fréquences/Volume	Coût (PPTTC, €)	Sources
Coût d'acquisition			
Kyprolis	4-6 inj sur 4 sem 24,1-26,1 mg/m ² 18 cycles max	■ €/mg ■ €/cycle modèle	Volume : Essai ASPIRE Prix : industriel
Velcade	4 inj sur 3 sem 1,18mg/m ² 8 cycles max	304,5€/mg 5680€/ cycle modèle	Volume : Essai ENDEAVOR Prix : BdM-IT, AMELI
Levact	120-150mg/ m ² Pendant 7 mois	2,80€/mg 1451€/ cycle modèle	Volume : HAS Prix : BdM-IT, AMELI
Imnovid	4mg/jour sur 21 jours Pendant 7 mois	433,93/UCD 9113€/ cycle modèle	Volume : RCP Prix : BdM-IT, AMELI
Revlimid	5-25mg/jour sur 21 jours Jusqu'à progression	Selon dosage	Volume : ENDEAVOR (Vd) ou ASPIRE (KRd et Rd) Prix : BdM-IT, AMELI
Dexamethasone	4 adm/cycle 40mg/jour 8 adm/cycle 20mg/jour Jusqu'à progression	0,27 €/mg 43-57€/cycle	Volume : RCP Prix : BdM-IT, AMELI
Coût d'administration			
Honoraire de dispensation	Revlimid, Innovid	22€ par cycle de traitement	
Injection	Velcade, Levact, Kyprolis	377€/séance	Valorisation : ENCC 2013
Transport	Velcade, Levact, Kyprolis	69€/séance	Touati, 2014
Coût de prise en charge			
Traitements concomitants		156 €	Etude AMGEN 2015
Hospitalisations (hors EI)		108 €	Etude AMGEN 2015
Consultations		23 €	Etude AMGEN 2015
Biologie		131 €	Etude AMGEN 2015
Imagerie		13 €	Etude AMGEN 2015
Coût des effets indésirables (par évènement)			
Neutropénie		7 138 €	Etude AMGEN 2015
Anémie		5 558 €	
Thrombocytopénie		6 945 €	Etude AMGEN 2015
Leucopénie		6 808 €	Borget et al, 2015
Neutropénie fébrile		6 808 €	Borget et al, 2015
Lymphopénie		6 808 €	Borget et al, 2015
Diarrhée		1 870 €	GHS 06M091 (Turini et al, 2015)
Fatigue		7 330 €	Etude AMGEN 2015
Constipation		1 870 €	GHS 06M091 (Turini et al, 2015)
Vomissement		1 870 €	GHS 06M091 (Turini et al, 2015)
Evènements thrombo-emboliques		5 703 €	Etude AMGEN 2015
Dyspnée		6 629 €	GHS 04M113 et 04M114
Hypophosphatémie		4 490 €	hypothèse (hypercalcémie)
Hypokaliémie			
Neuropathie pétéphérique		102 €	Etude AMGEN 2015
Pneumonie		5 657 €	Etude AMGEN 2015
Coût des soins palliatifs			

Coûts de traitement anti-tumoraux

Coûts d'acquisition

Les coûts associés aux traitements anti-tumoraux sont estimés à partir de la formule suivante :

Coût = (Nombre d'administration par cycle)*(Dose)*(Surface corporelle si applicable)*(prix du conditionnement adapté)/(volume par unité)

Avec : le nombre d'administration par cycle correspondant aux recommandations des RCP ;

la surface corporelle moyenne des patients est supposée de 1,73m² ;

les doses utilisées dans les essais ASPIRE et ENDEAVOR et non les doses fixées par le schéma d'administration théorique présenté dans les RCP. L'estimation des doses tient compte de l'inobservance sur la durée de traitement.

Des hypothèses de gaspillage sont formulées pour Kyprolis® (perte de produit de 6,5%) et pour Velcade® (perte de produit de 30%). Les différences de pertes entre les produits reflètent en particulier l'adaptabilité plus importante de Kyprolis® (flacons de 10, 20 et 30 mg).

Le coût des autres chimiothérapies utilisées dans le myélome multiple est inclus dans le montant du groupe homogène de séjour (GHS).

Pour les traitements prescrits après progression (IMNOVID®, LEVACT®), la durée de traitement utilisée dans le modèle était de 7 mois (durée de survie sans progression observée dans le registre des Hémopathies Malignes régional).

Autres coûts associés aux traitements anti-tumoraux

Le modèle comptabilise le coût de l'injection par voie parentérale, le coût de transport associé à l'injection des traitements et le coût de la dispensation des traitements sur la liste de rétrocession.

Le coût de l'administration par voie parentérale en HDJ est estimé à partir du tarif ENCC 2013 du GHS 28Z07Z (chimiothérapie pour tumeur en séance), duquel est soustrait les coûts des traitements facturés en sus et ajout des coûts de structure. Le coût estimé tient compte de la répartition public/ privé dans des proportions 85%/15%. Coût rapporté en Euro 2015.

Le coût de transport associé à une séance de chimiothérapie en HDJ a été extrait d'une publication (Touati et al., 2014). Cette étude a rapporté un coût de transport de 1€ par kilomètre en HDJ et une distance moyenne domicile-hôpital de 34km en moyenne. Les auteurs ont considéré un coût de 68,4€ par patient par injection.

Un forfait de 22€ est comptabilisé par cycle pour les produits délivrés en rétrocession (Revlimid, Imnovid).

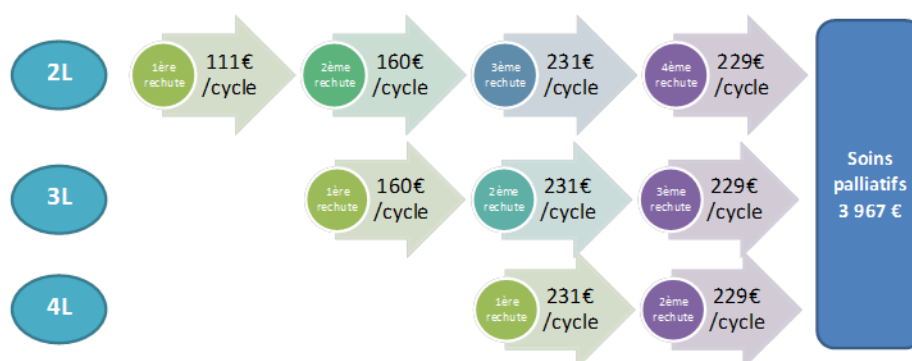
Coûts des évènements indésirables

Les coûts proviennent soit de l'étude de consommation de soin, soit de la littérature, soit des GHS spécifiques des effets indésirables

Coûts de soins non liés au traitement

Un coût de suivi non lié au traitement est appliqué par ligne de traitement, indépendamment du traitement reçu¹¹, jusqu'au passage à la ligne suivante.

¹¹ Le coût par cycle et par ligne de traitement est obtenu en pondérant les coûts observés pour chaque bras de traitement par leur effectif. En l'absence de données pour VELCADE® en 4^{ème} ligne, on fait l'hypothèse que les coûts sont similaires à ceux observés pour VELCADE® en 3^{ème} ligne.



Les coûts de prises en charge du patient non liés au traitement comprennent les coûts de soin ambulatoire (coûts des traitements concomitants, examens de laboratoires, frais d'imageries et frais de consultations) et les coûts d'hospitalisation (gestion de la maladie et soins palliatifs).

Tableau 11

Etude de consommation de soin Amgen	2 ^{ème} ligne	3 ^{ème} ligne	4 ^{ème} ligne	5 ^{ème} ligne
Coûts d'hospitalisation	143 €	47 €	143 €	216 €
Coûts ambulatoire	252 €	309 €	432 €	460 €
Coûts total de soin non lié au traitement	395 €	356 €	575 €	676 €

► Résultats de l'analyse de coût

Les coûts associés au traitement et à son administration sont moteurs dans le modèle. On observe une progression des coûts avec les lignes.

Quelle que soit la ligne de traitement, KRd est associé aux coûts de traitement les plus élevés. Rd représente plus de la moitié du coût d'acquisition dans le bras KRd en 2^{ème} et 3^{ème} lignes. La SSP de KRd étant plus courte au fur et à mesure des passages à la ligne supérieure, la durée de traitement par Rd seul dans le bras KRd était également plus courte. Du fait, la part de K dans le coût d'acquisition de KRd est augmentée en 3^{ème} ligne puis en 4^{ème} ligne.

Le coût associé à Rd dans le bras KRd est plus élevé que dans le bras Rd seul du fait des différences de SSP (tableau ci-dessus).

Vd est associé aux coûts des lignes de traitement post-progression les plus élevés, du fait d'une durée de survie post-progression plus longue que pour les autres comparateurs.

La diminution des SSP au fur et à mesure que l'on passe à une ligne de traitement supérieure est responsable d'une réduction des coûts d'acquisition des traitements et de transports, qui conduit à des coûts totaux inférieurs à la ligne précédente. Les temps de post-progression étant en revanche similaires entre les lignes, les coûts post-progression sont globalement similaires.

Tableau 12 Résultats de coûts, en euro, actualisés, par ligne de traitement

Traitements	Acq.	Trans.	Adm.	Hosp.	Ambu.	Eis	PP	Total
2^{ème} ligne								
KRd K Rd								280 499
Vd								150 423
Rd								148 313
3^{ème} ligne								
KRd K Rd								234 081
Vd								115 879
Rd								111 727
4^{ème} ligne								
KRd K Rd								195 185
Vd								85 254
Rd								86 291

PP : Post-Progression

Validation des simulations

L'industriel a comparé les coûts de suivi simulés avant progression par ligne (bras Rd) aux coûts observés dans l'étude de consommation des ressources. Les coûts simulés sont similaires aux coûts moyens estimés sur l'ensemble des patients de l'étude de consommation, mais ils sont différents des coûts observés uniquement chez les patients sous Revlimid®. Compte tenu du peu d'impact des coûts de suivi sur le RDCR, l'industriel considère que ces différences sont négligeables.

Tableau 13. Coûts moyens mensuels estimés avant progression

	Modélisation	Etude de coût (tous bras confondus)	Etude de coût (patients traités par REVLIMID®)
2^e ligne	429,4	428,0	616,46
3^e ligne	387,1	385,3	635,35
4^e ligne	625,1	623,1	550,71

L'industriel a comparé les durées moyennes de traitement simulées dans le modèle pour les bras KRd, Rd et Vd, aux durées observées sur données de vie réelle (référence non transmise). Il observe que les durées simulées sont plus importantes dans le modèle, particulièrement pour Rd. Selon l'industriel, la raison principale réside dans le fait que le lien entre arrêt de traitement et survie sans progression est directement issu de l'étude clinique dans le modèle.

Tableau 14. Durées moyennes de traitement

		KRd			Rd		Vd	
		K	R	D	R	D	V	D
2 ^{ème} ligne	Modèle	12,2	33,4	31,6	24	22,3	10,1	9,7
	Ref vie réelle		12		12		8	
3 ^{ème} ligne	Modèle	11,4	20,4	19,4	14,2	13,2	6,1	5,9
	Ref vie réelle		7		7		7	
4 ^{ème} ligne	Modèle	9,8	14,6	14,0	9,6	9,1	3,8	3,7

4.6.2 Analyse critique de l'évaluation des coûts

► Coûts pris en compte

Les coûts pris en compte sont conformes à la prise en charge du myélome multiple.

► Mesure, la valorisation et le calcul des coûts

Le principal problème dans l'évaluation des coûts tient aux durées de traitement qui ont un impact fort sur le RDCR. Il semble que les durées de traitement par Revlimid® soient surestimées dans le modèle, ce qui surestime les coûts de KRd et Rd. A contrario, le coût associé à Kyprolis® dans le bras KRd est sous-estimé par l'hypothèse d'une durée maximale de 18 cycles, vraisemblablement non réaliste.

4.7 Présentation des résultats et analyses de sensibilité

4.7.1 Présentation par les auteurs

► Résultats de l'étude économique

Dans l'analyse de référence, quelle que soit la ligne de traitement, le traitement KRd est le traitement le plus efficace parmi Rd, Vd et KRd. Il se situe toujours sur la frontière d'efficacité, Rd étant le traitement le précédant systématiquement sur la frontière d'efficacité.

KRd est associé à un RDCR versus Rd de 287 035 €/QALY, 292 960 €/QALY et 292 289 €/QALY pour la 2^{ème}, 3^{ème} et 4^{ème} ligne respectivement (tableau 11).

Les RDCR par ligne sont peu différents, ce qui met en évidence que l'efficacité est relativement conservée entre les lignes, même s'il existe une augmentation du RDCR avec l'augmentation du nombre de lignes antérieure. Le passage à une ligne supérieure se traduit par la baisse de la survie globale et par une baisse concomitante des coûts de traitements associés à la SSP. La réduction des coûts de traitement associée à la réduction de la SSP est insuffisante pour compenser la diminution du bénéfice clinique du KRd sur la SG.

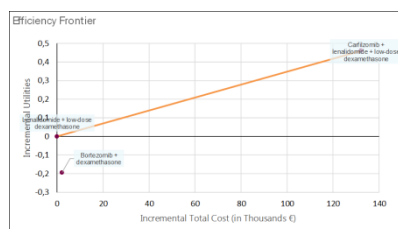
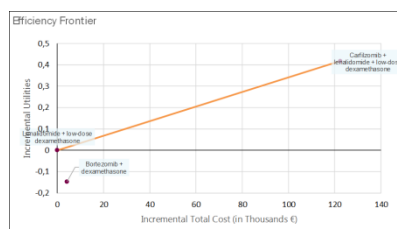
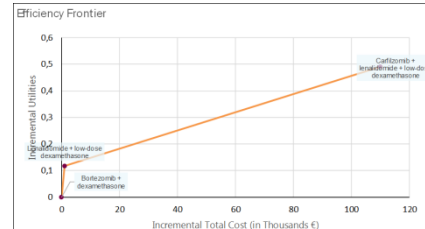
Tableau 15. RCDR de l'analyse de référence (QALY)

Traitements	QALY	Coûts (€)	ΔQALY	ΔCoûts (€)	RCDR (€ / QALY)
2^{ème} ligne					
Vd	2,51	150 423	-0,20	2 110	Dominée
Rd	2,71	148 313	0,00	0	Ref.
KRd	3,17	280 499	0,46	132 186	287 035
3^{ème} ligne					
Vd	1,68	115 879	-0,15	4 153	Dominée
Rd	1,83	111 727	0,00	0	Ref.
KRd	2,25	234 081	0,42	122 354	292 960
4^{ème} ligne					
Vd	1,24	85 254	0,00	0	Ref.
Rd	1,36	86 291	0,12	1 038	8 849
KRd	1,73	195 185	0,37	108 894	292 289

Figure 3 Frontières d'efficacité

Les frontières d'efficacité des 3 lignes de traitement sont très similaires.

On observe cependant que Vd fait partie de la frontière d'efficacité en 4^{ème} ligne, en raison de coûts totaux inférieurs à Rd.

2^{ème} ligne**3^{ème} ligne****4^{ème} ligne****► Prise en compte de l'incertitude et analyses de sensibilité****a) Incertitude liée aux choix structurants de l'évaluation****Horizon temporel**

Un horizon temporel de 25 ans a été considéré. L'industriel observe que la prise en compte d'un horizon temporel (HT) plus long a un impact limité sur le RCDR.

Le RCDR_{KRd vs Rd} est de 289 949 €/QALY en 2^{ème} ligne (-1%), de 289 889 €/QALY en 3^{ème} ligne (-1%) et de 289 949 €/QALY en 4^{ème} ligne (-1%).

Taux d'actualisation

Deux analyses étaient fournies, avec des taux d'actualisation de 0 et 2,5 %. Comme attendu, l'absence ou la réduction de l'actualisation réduit le RCDR. Le bénéfice de KRd en termes de survie étant accumulé tout au long de la simulation, l'actualisation se traduit par une valorisation réduite du bénéfice attendu sur le long terme. Au contraire, les coûts sont principalement accumulés durant les premiers cycles et sont moins sensibles à l'actualisation.

En 2^{ème} ligne, le RCDR_{KRd vs Rd} est de 253 833 €/QALY (-12%) avec un taux d'actualisation de 0%, et de 274 550 €/QALY (-4%) avec un taux d'actualisation de 2,5%.

En 3^{ème} ligne, le RCDR_{KRd vs Rd} est de 264 220 €/QALY (-10%) avec un taux d'actualisation de 0%, et de 282 251 €/QALY (-4%) avec un taux d'actualisation de 2,5%.

En 4^{ième} ligne, le RDCR_{KRd vs Rd} est de 265 597 €/QALY (-9%) avec un taux d'actualisation de 0%, et de 282 396 €/QALY (-3%) avec un taux d'actualisation de 2,5%.

b) Incertitude liée aux choix de modélisation (hypothèses et sources de données)

Impact des hypothèses sur la modélisation de l'efficacité de Rd

- ❖ *Impact du choix d'utiliser les données issues du registre français des hémopathies malignes pour modéliser les courbes de survie SSP et SG associées au traitement Rd*

Deux analyses de sensibilité ont été réalisées pour appréhender l'impact du choix des données du registre pour estimer la SG et SSP du bras Rd.

- Si les courbes de survie associées à Rd sont modélisées à partir des données issues du bras Rd de l'essai ASPIRE, l'analyse de sensibilité montre que le RDCR est augmenté.

Le RDCR_{KRd vs Rd} est de 317 336 €/QALY en 2^{ième} ligne (+11%), de 338 787 €/QALY en 3^{ième} ligne (+16%) et de 371 921 €/QALY en 4^{ième} ligne (+27%).

Selon l'industriel, cette augmentation par rapport à l'analyse de base est principalement due à des durées de SSP et de SG plus faibles dans le registre que dans l'essai ASPIRE (SSP médiane¹² : 16,6 mois vs 17,6 mois ; SG médiane : 35,6 mois vs 40,5 mois).

- Les courbes de survie ont été estimées sur la population totale du registre. Si les courbes de survie associées à Rd sont modélisées en utilisant la borne supérieure de l'estimation de la SG et de la SSP, qui s'approche des données observées en ne retenant que les patients du registre sous Rd, l'analyse de sensibilité montre que le RDCR est augmenté.

Le RDCR_{KRd vs Rd} est de 311 947 €/QALY en 2^{ième} ligne (+9%), de 325 451 €/QALY en 3^{ième} ligne (+11%) et de 325 451 €/QALY en 4^{ième} ligne (+1%).

L'industriel note que les résultats de cette analyse de sensibilité sont relativement similaires aux résultats de l'analyse de référence, bien que plus élevés.

- ❖ *Impact du choix d'une fonction paramétrique de Weibull pour extrapoler les données de SSP et de SG issues du registre français des hémopathies malignes*

Deux analyses de sensibilité déterministes testaient la fonction de survie paramétrique de la SSP et SG utilisées pour le bras Rd.

- Lorsque la SSP et la SG sont modélisés par une distribution log-logistique (meilleur ajustement sur les critères AIC et BIC en ligne 3 et ligne 4), le RDCR_{KRd vs Rd} est de 251 083 €/QALY en 2^{ième} ligne (-13%), de 204 312 €/QALY en 3^{ième} ligne (-30%) et de 235 317 €/QALY en 4^{ième} ligne (-19%).
- Lorsque la SSP et la SG sont modélisés par une distribution exponentielle (moins bon ajustement sur les critères AIC et BIC), le RDCR_{KRd vs Rd} est de 311 018 €/QALY en 2^{ième} ligne (+8%), de 320 914 €/QALY en 3^{ième} ligne (+10%) et de 342 910 €/QALY en 4^{ième} ligne (+17%).

Impact des hypothèses sur la modélisation de l'efficacité de Kyprolis® et Vd

- ❖ *Impact de l'hypothèse d'une durée d'efficacité de Kyprolis® sur 36 mois*

Les auteurs ont également conduit 2 analyses de sensibilité déterministes sur la durée d'efficacité de Kyprolis®, en supposant qu'elle dure 32,3 mois et sur la vie entière

- Efficacité de Kyprolis® sur vie entière : le RDCR_{KRd vs Rd} est de 182 708 €/QALY en 2^{ième} ligne (-36%), de 196 354 €/QALY en 3^{ième} ligne (-33%) et de 207 783 €/QALY en 4^{ième} ligne (-29%).
- Efficacité de Kyprolis® sur 32,3 mois (médiane de suivi pour la SG) : le RDCR KRd vs Rd étaient de 303 353 €/QALY en 2^{ième} ligne (+6%), de 307 840 €/QALY en 3^{ième} ligne (+5%) et de 305 454 €/QALY en 4^{ième} ligne (+5%).

¹² Données extrapolées à l'aide de la fonction exponentielle, médiane non atteinte dans l'essai.

❖ *Impact de l'hypothèse d'une efficacité de Vd modélisée en appliquant un HR constant sur 18 mois (0,521 pour la SSP ; 0,884 pour la SG)*

Deux analyses de sensibilité déterministes étaient présentées intégrant des HR entre Rd et Vd variables dans le temps, issus d'une analyse de cox par morceaux [HR de 1,0261 (SG) et 0,5759 (SSP) de 0 à 12 mois, puis de 0,6103 (SG) et 0,3423 (SSP) après 12 mois].

Les deux analyses montrent que cela n'a pas d'impact sur la frontière d'efficacité, avec des RDCR toujours estimés versus Rd. Cette hypothèse n'a donc aucun impact sur le RDCR de KRd vs Rd.

Impact de la méthode de prise en compte des événements indésirables

- Les méthodes alternatives testées (absence de comptabilisation des événements indésirables ou prise en compte sur le premier cycle) n'a pas d'impact sur le RDCR.
- Afin de pouvoir discuter de l'impact de potentielles sous-estimations des effets secondaires aux traitements, une analyse de sensibilité ajoutant des désutilités en lien avec chacun des effets indésirables a été réalisée (cf. hypothèse de l'analyse de référence supposant que les valeurs d'utilité issues du questionnaire QLQ-C30 captaient l'impact sur la qualité de vie des effets indésirables et des hospitalisations). Cette analyse n'a montré aucun impact sur le RDCR.

Choix des méthodes et des sources de données pour documenter les valeurs d'utilité

❖ *Impact du choix de la source de données*

- Source 1 : valeurs utilités transformées à partir des scores QLQ-C30 de l'essai ASPIRE selon l'algorithme de Longworth et al. Le RDCR_{KRd vs Rd} est de 294 235 €/QALY en 2^{ème} ligne (+3%), de 307 758 €/QALY en 3^{ème} ligne (+5%) et de 311 518 €/QALY en 4^{ème} ligne (+7%).
- Source 2 : valeurs d'utilité obtenues dans l'essai MM003 avec application d'une désutilité associée aux événements indésirables (essai incluant des patients plus sévères avec un nombre plus important de traitements antérieurs). Le RDCR_{Rd vs Rd} est de 280 692 €/QALY en 2^{ème} ligne (-2%), de 286 787 €/QALY en 3^{ème} ligne (-2%) et de 283 231 €/QALY en 4^{ème} ligne (-3%).
- Source 3 : valeurs d'utilité de van Agthoven 2004 avec application d'une désutilité associée aux événements indésirables. Le RDCR_{KRd vs Rd} est de 314 937 €/QALY en 2^{ème} ligne (+10%), de 284 199 €/QALY en 3^{ème} ligne (-3%) et de 260 594 €/QALY en 4^{ème} ligne (-11%).

❖ *Impact de la méthode d'estimation des valeurs d'utilité*

- Le modèle repose sur une méthode fondée sur l'estimation de l'utilité en L3. Une désutilité associée à la progression (-0.050) est soustraite pour estimer l'utilité en L2 ; elle est ajoutée pour estimer l'utilité en L4. Une analyse de sensibilité a été réalisée en n'intégrant qu'une valeur d'utilité, indépendamment de la ligne de traitement. Le RDCR_{KRd vs Rd} est de 322 926 €/QALY en 2^{ème} ligne (+13%), de 295 394 €/QALY en 3^{ème} ligne (+1%) et de 273 480 €/QALY en 4^{ème} ligne (-6%).

Données toutes lignes de traitement confondues (essai ASPIRE)					
Pré-progression			Post-progression		
KRd	Rd	Vd	KRd	Rd	Vd
0,7390	0,7220	0,7220	0,6814	0,6814	0,6814

- Une analyse de sensibilité a été réalisée permettant d'accroître la désutilité associée au passage d'une ligne 4 à une ligne 5, en ajoutant une désutilité de 0,17. Le RDCR_{KRd vs Rd} est de 269 407 €/QALY en 2^{ème} ligne (-8%), de 287 279 €/QALY en 3^{ème} ligne (-2%) et de 290 745 €/QALY en 4^{ème} ligne (-1%).

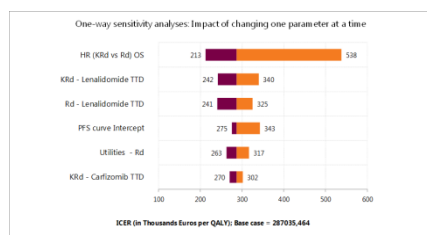
c) Incertitude liée à la variabilité des paramètres du modèle

Analyse de sensibilité déterministe

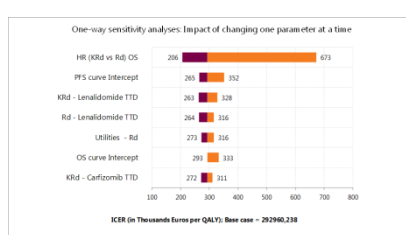
❖ *Principales variables ayant un impact sur le RDCR_{KRd vs Rd}*

Seules les variables ayant le plus d'impact sur le RDCR_{KRd vs Rd} ont été reprises dans le rapport technique transmis par l'industriel, sous la forme d'un graphique de tornado.

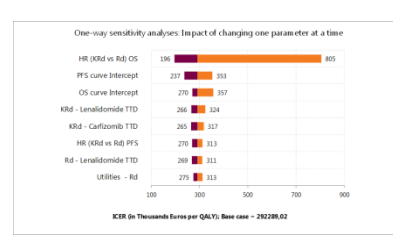
2^{ème} ligne



3^{ème} ligne



4^{ème} ligne



Les paramètres influençant le plus le RDCR_{KRd vs Rd} sont tous associés :

- À l'extrapolation des données de survie : HR (OS) pour KRd vs Rd et les paramètres d'extrapolation de la survie.
- À la durée des traitements, en particulier la durée de traitement sous Rd (bras Rd et KRd) et sous Kyprolis® (bras KRd).

Ces résultats sont liés à l'incertitude qui existe autour de ces paramètres du fait d'une part de la relative immaturité des données APIRE (HR et durées de traitement) et d'autre part à la faible taille d'échantillon des données de registre (extrapolation de la survie).

❖ Analyse spécifique sur l'impact d'une variation du prix de Kyprolis® sur le RDCR.

L'analyse de sensibilité réalisée sur le prix montre que la variation du RDCR est moins que proportionnelle à la variation du prix.

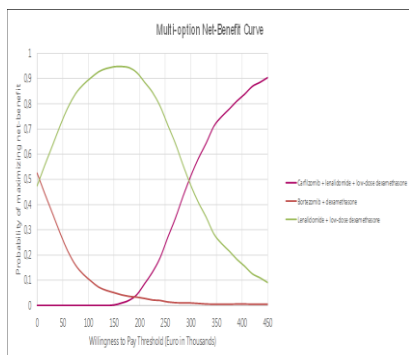
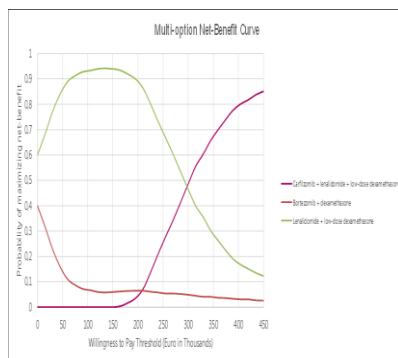
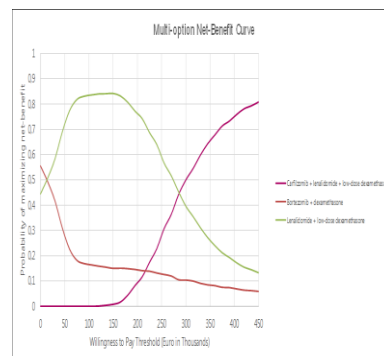
- Une réduction de 50% du prix de Kyprolis® réduit le RDCR_{KRd vs Rd} de 29% en 2^{ème} ligne (205 108 €/QALY), de 29% en 3^{ème} ligne (208 515 €/QALY), de 28% en 4^{ème} ligne (209 672€/QALY).
- Une augmentation de 20% du prix de Kyprolis® augmente le RDCR_{KRd vs Rd} de 11% en 2^{ème} ligne (319 806 €/QALY), de 12% en 3^{ème} ligne (326 738 €/QALY), de 11% en 4^{ème} ligne (325 336 €/QALY).

Analyse de sensibilité probabiliste

L'analyse de sensibilité probabiliste (PSA) teste l'impact sur le RDCR_{KRd vs Rd} d'une variation aléatoire combinée des données de survie, des données de coût et d'utilité, introduites dans le modèle.

Elle a été réalisée par une simulation de deuxième ordre de Monte Carlo sur 1000 simulations en utilisant l'approche du bénéfice net. Pour chacun des paramètres testés, l'industriel a présenté la valeur de référence du paramètre, la loi de distribution ainsi que les paramètres de la loi (cf. rapport technique, mars 2016).

Le RDCR estimé dans l'analyse de base Kyprolis® correspond à la disposition à payer à partir de laquelle Kyprolis® présente la plus forte probabilité de maximiser le bénéfice net sur 1000 simulations. En termes d'incertitude, l'analyse probabiliste confirme qu'au niveau des RDCR estimés dans l'analyse de base, le choix de favoriser Kyprolis® représente un risque de 50% de choisir le produit qui ne maximise pas le bénéfice net. Cependant, on observe que le risque se réduit fortement sur des dispositions à payer plus élevées, ce qui indique que l'incertitude générée par la variabilité aléatoire des valeurs entrées dans le modèle est faible.

2^{ème} ligne

3^{ème} ligne

4^{ème} ligne


► Analyses de scénario

L'industriel a présenté deux analyses en scénario, à la demande de la HAS.

Tableau 16 : Hypothèses introduites dans les analyses en scénario

	Analyse de scénario optimiste	Analyse de référence	Analyse de scénario NMA
Comparateurs	2L : Rd, Vd 3L : Rd, Vd 4L : Rd, Vd	2L : Rd, Vd 3L : Rd, Vd 4L : Rd, Vd	2L : Rd, Vd, Td 3L : Rd, Vd, Benda, Pom-d, Td, 4L : Rd, Vd, Benda, Pom-d, Td,
Durée efficacité KRd par rapport à Rd	Vie entière	36 mois	36 mois
Fonction paramétrique SG et SSP de Rd	Log logistique	Weibull	Weibull
Gaspillage	Non	Oui (6.5%)	Oui (6.5%)
Horizon temporel	Vie entière	15 ans	15 ans
Durée de simulation	25 ans	15 ans	15 ans

Résultats
Tableau 17 : Résultats des analyses en scénario optimiste

Traitements	Scénario optimiste	Analyse de référence
2^{ème} ligne		
Vd	Ref	Dominée
Rd	27 821	Ref.
KRd	143 468	287 035
3^{ème} ligne		
Vd	Dominée	Dominée
Rd	Ref	Ref.
KRd	141 607	292 960
4^{ème} ligne		
Vd	Ref	Ref.
Rd	17 341	8 849
KRd	106 686	292 289

❖ Scénario optimiste

Le scénario optimiste permet d'estimer les $RDCR_{KRd \text{ vs } Rd}$, si on considère une augmentation des différences de survies entre KRd et Rd par rapport à ce qui est considéré dans l'analyse de référence (choix de la fonction d'extrapolation et application des HR de KRd versus Vd sur la vie entière). Associé à l'absence de gaspillage de flacon (réduction des coûts de traitement, les RDCR sont plus faibles. Par ailleurs, contrairement à l'analyse de référence, Vd est sur la frontière d'efficience en 2^{ème} et 3^{ème} lignes.

Globalement les résultats des analyses de sensibilité sur ce scénario sont similaires à l'analyse de référence en ce qui concerne les hypothèses structurantes et les analyses déterministes.

Les analyses de sensibilité probabilistes reflètent l'inclusion de Vd sur la frontière d'efficience. Quelle que soit la ligne, la courbe se décompose en trois parties :

- En dessous de 40 000 € / QALY, Vd est le traitement qui maximise le bénéfice net.
- Entre 40 000 et 150 000 € / QALY, Rd est le traitement qui maximise le bénéfice net
- Au-dessus de 150 000 € / QALY, KRd est le traitement qui maximise le bénéfice net

► Validité et transférabilité des résultats

L'industriel discute les résultats de son évaluation au regard des analyses de sensibilité réalisées, pour identifier les principales sources d'incertitude.

L'utilisation des données de registre permet, selon l'industriel, de garantir la comparabilité de la population simulée à la population cible et la transférabilité des résultats en vraie vie. Il note cependant que l'utilisation des données tous traitements confondus est une limite, mais dont l'impact est peu important sur le RDCR (RDCR plus élevé).

La méthode utilisée pour estimer des utilités sans progression croissantes avec la ligne d'instauration de traitement conduit également à des RDCR plus faibles que si l'on considérait une utilité sans progression identique dans les trois sous-populations, surtout dans la sous-population avec une instauration en 2^{ème} ligne qui bénéficie d'un gain d'utilité lié au gain de SSP plus important. L'algorithme de mapping retenu conduit également à des RDCR plus faibles que l'option alternative, laquelle aurait cependant diminué la transposabilité des résultats en raison des spécificités de la population sur laquelle il a été développé.

Outre ces éléments, l'hypothèse concernant une durée limitée de l'effet traitement de Kyprolis® dans le temps est une source importante d'incertitude, défavorable au produit. Selon l'industriel, l'hypothèse selon laquelle les bénéfices sur la survie observés à court/moyens termes sont maintenus à long terme est étayée par l'avis de cliniciens, ainsi que par de récents essais cliniques dans le myélome multiple (VELCADE®^{33,67} ; REVLIMID®^{32,37,39}) et également des analyses en vie réelle^{50,68}. L'hypothèse d'un effet traitement de KRd sur Rd maintenu sur la vie entière réduit considérablement le RDCR.

Toutefois, en cas de poursuite du traitement jusqu'à progression, ces gains de survies sont compensés par l'augmentation des coûts de traitement, qui augmenterait les RDCR par rapport à l'analyse de référence. Cette augmentation des coûts s'explique en grande partie par le coût de Rd qui représente plus de la moitié des coûts de traitement de 2^{ème} ligne. Ceci explique également pourquoi la variation du RDCR n'est pas proportionnelle à la variation du coût de Carfilzomib.

Enfin, les analyses déterministes montrent que l'incertitude autour de l'efficacité de KRd sur la survie globale, la durée de traitement de KRd et Rd et l'extrapolation des données de survie ont un impact non négligeable sur le RDCR. Concernant le premier point, les données d'efficacité de KRd sont issues de l'étude ASPIRE dans laquelle la médiane de survie globale n'a pas été atteinte à l'issue du suivi des patients. De fait, cela conduit à un intervalle de confiance du HR KRd vs Rd relativement large. Concernant le second point, les coûts de traitements, proportionnels aux durées de traitements, représentent le poste de coût le plus important dans l'analyse.

4.7.2 Analyse critique de la présentation des résultats et de l'analyse de sensibilité

► Résultats de l'étude économique

Les résultats de l'analyse de référence et des analyses de scénario sont présentés de manière claire et adaptée.

► Analyse de l'incertitude

Incertitude structurelle (hypothèses et choix méthodologiques de modélisation)

L'horizon temporel de 25 ans testé en analyse de sensibilité n'apporte rien, puisque le pourcentage de patients vivants à 15 est faible. Il aurait été préférable de tester une durée de simulation plus courte (entre 10 ans et 15 ans).

Les principaux choix méthodologiques qui vont avoir une influence sur le résultat ont été correctement identifiés par l'industriel. Cependant, l'interprétation proposée est critiquable pour trois d'entre eux.

- L'utilisation des données issues du registre d'hématologie maligne pour estimer les courbes de survie sous Rd. S'il est exact que la population du registre est plus représentative d'une population réelle que la population de l'essai, il reste deux critiques importantes qui correspondent à des choix favorables au produit. Premièrement, les données utilisées correspondent au TCM et non au traitement Rd, avec un impact favorable au produit. Deuxièmement, le $HR_{KRd \text{ vs } Rd}$ appliqué sur cette population a été estimé sur une population présentant un pronostic moins dégradé, ce qui conduit de nouveau à favoriser le produit.
- L'utilisation d'un algorithme de mapping sur l'utilité favorable au produit. S'il est vrai que la population sur laquelle l'algorithme retenu a été développé est plus proche de la population de l'essai, l'utilisation de l'algorithme alternatif aurait permis de corriger un peu la qualité de vie observée dans l'essai, vraisemblablement surestimée par rapport à ce que l'on peut attendre en vie réelle.
- La durée de l'effet traitement, qui pourrait être sous-estimée dans le modèle. Il est à noter, que même en retenant l'hypothèse la plus favorable au produit, à savoir une durée d'effet vie entière, les RDCR estimés restent encore compris entre 183K €/QALY et 208K €/QALY. Ces RDCR sont estimés avec une durée de traitement maximale de 18 cycles pour Kyprolis®, ce qui laisse présager des RDCR avec des durées de traitement plus réalistes encore plus élevés.

Incertitude sur la variabilité des valeurs des paramètres

Sous les hypothèses et les choix méthodologiques retenus dans l'analyse de référence, l'incertitude associée à la variabilité aléatoire des données entrées dans le modèle peut être considérée comme faible.

Annexe 3 – Faisabilité de la méta-analyse

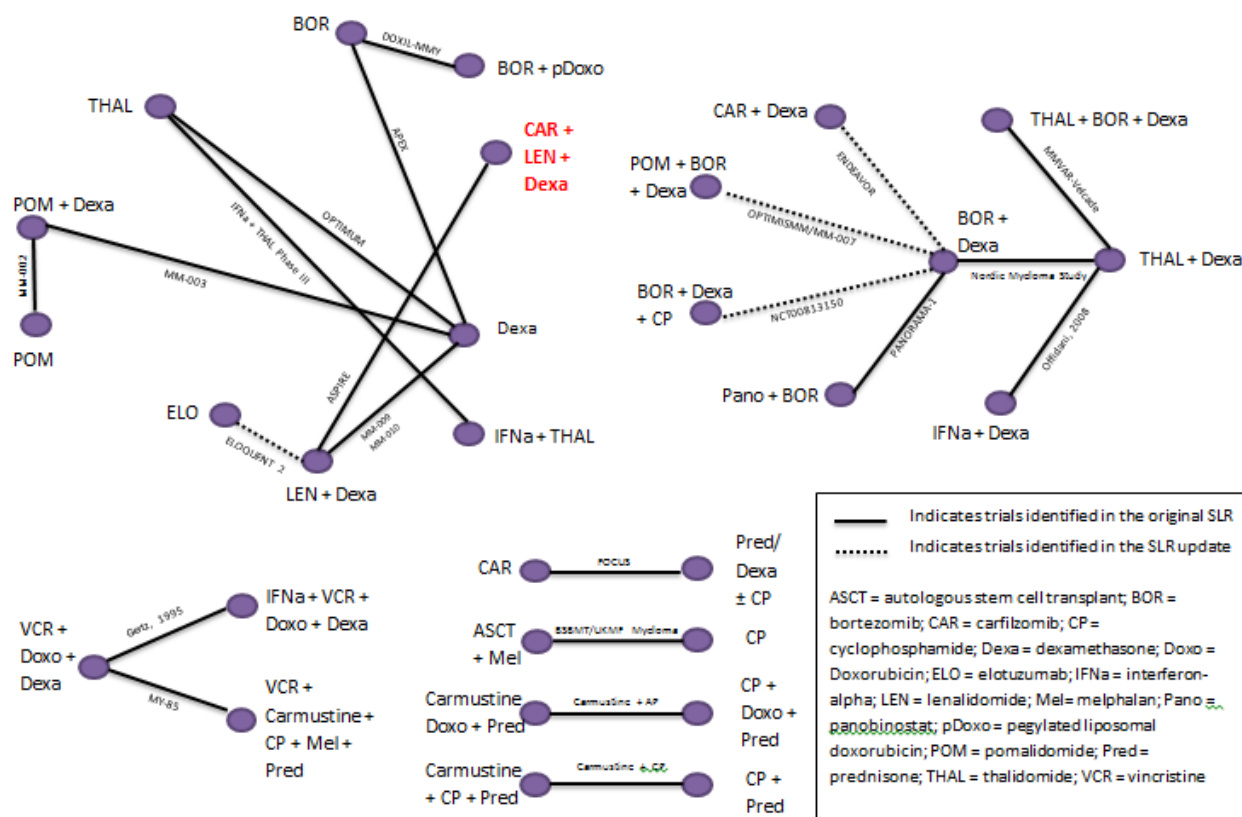
Telle que présentée par les auteurs

Méthode

Une revue systématique de la littérature a été réalisée en juillet 2015 de manière à identifier l'ensemble des essais cliniques disponibles dans le myélome multiple en rechute. Les traitements sélectionnés étaient les suivants : les associations à base de KYPROLIS®, VELCADE®, REVLIMID®, thalidomide, IMNOVID®, LEVACT®, FARYDAK®, EMLICITI®, autres chimiothérapies, placebo et traitement de fond optimisé. L'ixazomib n'a pas été inclus par les auteurs car les données publiées n'étaient pas disponibles au moment de la revue. Une étude de faisabilité de cette méta-analyse a été fournie par les auteurs en annexe du dossier. Les critères de résultats sélectionnés étaient un HR sur l'analyse de la SG et de la SSP entre les deux traitements comparés ainsi qu'une mesure de dispersion ou la p-value du test étaient rapportés.

23 études ont été retenues pour une inclusion potentielle dans une méta-analyse en réseau. Trois réseaux d'études ont ainsi été construits par les auteurs (Figure 4). Le réseau KRd inclut Rd, IMNOVID® en association avec la dexaméthasone, VELCADE® en monothérapie ou en association avec la doxorubicin et EMLICITI® en association à REVLIMID® et à la dexaméthasone (ERd). Le détail complet de la revue systématique de littérature est fourni dans une autre annexe du rapport technique.

Figure 4. Réseaux des études issues de la revue systématique de la littérature



Les auteurs justifient la non-faisabilité d'une méta analyse en réseau dans leur dossier avec les arguments suivant :

- Krd n'est pas connecté à certains comparateurs (Vd, Td, PanoVd)
- Des doses différentes de dexaméthasone sont utilisées dans les essais : Faible doses pour ASPIRE et ELOQUENT2 ; fortes doses pour les études MM-009 et MM-010. Les auteurs reportent une étude chez les patients avec un myélome multiple nouvellement diagnostiqué que l'efficacité de REVLIMID® associé avec une faible dose de dexaméthasone (Rd) ou une forte dose de dexaméthasone (RD) était différente (survie sans progression de 25.3 mois pour le bras Rd et de 19.1 mois pour le bras RD, $p=0.026$)⁴⁰.
- Les populations des essais sont hétérogènes sur certains critères :
 - o Nombre de traitements précédemment reçus :
 - Entre 1 et 3: 4 études (APEX, ASPIRE, ELOQUENT-2, ENDEAVOR, OPTIMUM, PANORAMA-1)
 - ≥ 1 : 3 études (MM-009, MM-010, DOXIL-MMY-3001)
 - ≥ 2 : 2 études (MM-002, MM-003)
 - o Type de traitements précédents (ex : lénalidomide est un critère d'exclusion pour MM009 et MM010, et d'inclusion pour MM002 et MM003)
 - o Antécédents d'autogreffe (ex : critère d'inclusion dans MMVAR et pas dans les autre études)
 - o Niveau de risque cytogénétique Les critères d'inclusions sur le nombre de traitements précédents varient de la manière suivante

Les auteurs reportent que les limites méthodologiques importantes de méta-analyse dans le myélome ont été également reportées dans des publications antérieures (Ruggeri et al. 2015). D'autre part, les auteurs soulignent que même si la faisabilité de la méta-analyse est forcée avec des hypothèses fortes, le réseau d'étude utilisé serait relativement restreint au regard du nombre de comparateurs inclus et que la majorité des comparaisons ne reposerait que sur une étude. Cela impliquerait que la sensibilité de la méta- analyse à l'inclusion ou non d'une étude ne pourrait être testée.

Analyse critique du SEESP

La revue systématique de la littérature menée par les auteurs est de bonne qualité méthodologique. La qualité méthodologique des essais a été évaluée suivant les critères recommandés par le NICE et la Cochrane, en intégrant les mesures du risque de biais.

Cette revue systématique met en évidence le fait que les traitements ne peuvent pas être connectés au sein d'un réseau sans recourir à des hypothèses cliniques fortes.

Les arguments avancés par les auteurs sur la fragilité de la méta-analyse sont réalistes, à savoir le caractère fort des hypothèses prises et l'hétérogénéité des populations comparées.

D'autre part, il est important de noter la fragilité du réseau si il était créé, car pour chaque comparaison directe, un seul essai est généralement disponible. Le niveau de preuve atteint ne serait pas celui d'une méta-analyse conventionnelle, ni celui d'une méta-analyse en réseau basée sur un réseau solide. Un réseau solide serait composé des comparaisons directes, incluant plusieurs essais, et des boucles fermées, permettant de tester l'hypothèse d'incohérence).

Le SEESP questionne la faisabilité d'une méthode MAIC combinant données agrégées et données individuelles en réajustant pour obtenir des pseudo-populations comparables sur ces variables (Signorovitch 2010).

A ce jour, la quantification du niveau de preuves d'une comparaison indirecte via une analyse poolée de 2 essais avec ajustement sur des variables individuelles n'est pas évidente. Il s'agit d'une comparaison indirecte via une analyse poolée de 2 essais avec ajustement sur des variables individuelles. Cette méthode est sujette à fournir des estimations biaisées. Une méta-analyse en réseau sur données individuelles aura un niveau de preuve supérieur à la MAIC car la méta-analyse en réseau permet de conserver les bénéfices liés à la randomisation. Des travaux méthodologiques comparant ces 2 méthodes sur des bases de données individuelles (réelles ou simulées) permettraient de mettre en évidence les différences de biais (quantifier) et de fournir un guide pour l'utilisation de ces méthodes.

Annexe 4 – Echange avec l'industriel

La liste de questions techniques ci-dessous a été adressée à l'industriel ; elle est commune aux deux dossiers évalués simultanément.

L'industriel a adressé des réponses écrites à la HAS.

Les éléments en gras doivent être traités en priorité. Lorsque des modifications de la modélisation sont demandées, l'ensemble des analyses mises à jour doit être fourni (y compris les analyses de sensibilité).

Questions diverses

- 1) Le dossier de présentation et le rapport technique présentent des informations contradictoires concernant le statut de médicament orphelin de Kyprolis®. Merci de corriger.
- 2) Indiquer, de manière synthétique, les essais de phase II/III actuellement en cours dans le myélome multiple en rechute, ainsi que les essais incluant carfilzomib, financés par Amgen, actuellement en cours dans d'autres indications (libellé de l'indication, nombre d'essais, dates estimées des résultats).

Analyse de référence

- 3) **Il est attendu que l'objectif de l'évaluation médico-économique soit précisé ou, le cas échéant, redéfini au regard des comparateurs pertinents selon la ligne de traitement, et des données d'efficacité spécifique à chaque ligne.**
Il est attendu que l'analyse de référence repose sur des choix structurants cohérents avec l'objectif ainsi redéfini (cf. Q5, Q6, Q7).

Explication de la question.

L'objectif est défini par l'industriel sur l'ensemble de l'indication du produit, c'est-à-dire dès la première rechute et sans distinguer les lignes d'instauration. Or, d'une part, l'agressivité de la maladie augmente à chaque rechute comme le montre l'étude AMGEN 2015 (réduction des TTP selon la ligne de traitement) et, d'autre part, les traitements utilisés en pratique clinique sont différents selon les lignes.

Il existe donc une distorsion dans le modèle déposé entre :

- *les données d'efficacité, issues de patients préalablement traités toutes lignes confondues (43% en 2^{ème} ligne et 57% en 3/4^{ème} ligne), et*
- *le choix des comparateurs (deuxième ligne uniquement).*

C'est pourquoi, le modèle actuel ne permet pas de démontrer de manière adaptée l'efficience de Kyprolis®. Son interprétation en termes d'efficience toutes lignes confondues se heurte à la difficulté de comparateurs hétérogènes selon les lignes d'instauration. Son interprétation en termes d'efficience d'une instauration en seconde ligne de traitement se heurte à l'hétérogénéité des données d'efficacité entrées dans le modèle.

Il est noté que l'argument présenté pour exclure certains traitements (« la méta-analyse en réseau ne peut être exécutée en raison de l'hétérogénéité des populations des essais ») confirme l'intérêt d'identifier des sous-populations par ligne d'instauration de traitement.

Il est attendu que l'industriel explore la possibilité de présenter une analyse par sous-population, permettant de prendre en compte de manière adaptée les comparateurs pertinents. Une discussion sur la robustesse d'une telle analyse sur le plan statistique devrait orienter l'industriel sur le choix de la présenter en analyse de référence ou en analyse complémentaire.

Dans le cas où cette analyse en sous-population ne serait pas considérée par l'industriel comme possible, il est attendu qu'il redéfinisse l'objectif de l'évaluation présentée en cohérence avec les comparateurs retenus (comparateurs de seconde ligne) en explorant l'incertitude associée avec les données entrées dans le modèle, observées en seconde ligne et plus. Dans ce cas, l'industriel devra estimer la part que représente cette population d'analyse restreinte sur la population cible de remboursement.

Analyses en scénario

- 4) Au vu de l'incertitude générée par l'immaturité des données cliniques de l'essai ENDEAVOR et des conditions d'utilisation de carfilzomib en vie réelle, il serait intéressant de pouvoir disposer en plus de l'analyse de référence, d'une analyse en scénario « optimiste » et « pessimiste ». Le tableau de synthèse ci-dessous reprend les hypothèses spécifiques aux trois analyses attendues.

Les explications pour les changements attendus dans l'analyse de référence sont développées dans la liste de questions ci-dessous.

Tableau récapitulatif

	Analyse de scénario Optimiste	Analyse de référence	Analyse de scénario pessimiste
Durée efficacité Krd par rapport à Rd	Vie entière	36 mois	32,3 mois
Fonction paramétrique survie pour la SG de Rd	Exponentielle	Weibull (Q12)	Weibull
Gaspillage	Non (ou maintien de l'hypothèse actuelle)	Oui (Q17)	Oui
HT	Vie entière	15 ans (Q10)	15 ans
Durée de simulation	25 ans	15 ans	15 ans

- 5) Il est attendu que soit présentée une analyse de sensibilité sur la méta-analyse en forçant sa faisabilité, à partir d'hypothèses cliniques fortes (p.22 Annexe 6) :
- une hypothèse d'équivalence d'efficacité entre les faibles et fortes doses de dexaméthasone, permettant d'inclure bortezomib dans la méta-analyse ;
 - une hypothèse d'équivalence d'efficacité entre BOR et BOR+Dexa de préférence (et/ou THAL et THAL+Dexa) permettant de relier les deux réseaux déconnectés.

Il est attendu une analyse en scénario de l'évaluation économique intégrant les résultats de cette nouvelle méta-analyse. Si les moyens dont disposent les auteurs et le temps imparti lors de cet échange technique ne permettent pas de mettre en œuvre une approche Bayésienne, le recours à une approche fréquentiste en utilisant le package Netmeta du logiciel R [¹³], développé par G. Rücker [¹⁴], peut être envisagée.

Les analyses de sensibilités sont attendues sur cette analyse en scénario.

Choix structurants de l'évaluation

Population d'analyse

- 6) **Il est attendu que l'industriel définisse sa population d'analyse en fonction des éléments développés ci-dessus (Q3).**

Explication de la question : la définition d'une population d'analyse agrégée sans sous-

¹³ <https://cran.r-project.org/web/packages/netmeta/netmeta.pdf>

¹⁴ Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods. 2012 Dec;3(4):312-24.

population n'est pas valide, si les comparateurs par ligne de traitement ne sont pas les mêmes. En fonction de l'objectif fixé en Q3, l'industriel pourra définir la population d'analyse uniquement en deuxième ligne ou sur l'ensemble des patients pour lesquels un remboursement avec ASMR II est revendiqué, si des sous-populations d'analyse sont définies pour prendre en compte l'hétérogénéité des comparateurs en fonction de la ligne d'instauration de traitement.

- 7) Pouvez-vous préciser la place de votre produit dans la prise en charge des patients éligibles à une autogreffe, ainsi que l'impact attendu de l'introduction de cette trithérapie chez ces patients. Justifier l'absence de sous-population d'analyse distinguant ces patients. Le cas échéant, une sous-population d'analyse pourra être définie.

Explication de la question : l'étude AMGEN 2015 (focus study) montre que le statut SCT influence le traitement, en particulier avec des durées de traitement plus courtes et des durées sans traitement plus longues. Il est donc attendu que les coûts soient différents selon le statut SCT. Elle montre également que les patients SCT en 2nde ligne progressent plus tard (bien que la différence soit moins marquée qu'en 1^{ère} ligne) et qu'ils ont une survie globale plus longue.

Comparateurs

- 8) **Il est attendu que la sélection des comparateurs inclus dans l'analyse de référence soit revue conformément aux éléments développés ci-dessus (Q3).**

Explication de la question : les comparateurs retenus dans la version initiale du modèle correspondent aux options thérapeutiques de 2nde ligne. Aucun comparateur de 3^{ème} ligne et plus n'est retenu.

- 9) **Il est attendu que Kd soit inclus dans l'analyse médico-économique.**

Explication de la question : Compte tenu du recours à l'essai de phase III Endeavor dans l'analyse groupée afin d'intégrer Vd, il est attendu que les données d'efficacité de Kd soient présentées et Kd comparés à KRd dans l'analyse coût-efficacité. D'autre part, Kd est déjà intégré dans des recommandations récemment publiées (cf. NCCN Clinical Practice Guidelines in Oncology, version 2.2016).

- 10) Connaissez-vous la proportion de bortézomib prescrit ou non en association à la dexaméthasone ? (Annexe 2. Figure 14)

Horizon temporel

- 11) Au vu des données épidémiologiques en France, notamment de l'âge des patients préalablement traités (>70 ans) et des médianes de survie des myélomes multiples, il semble qu'un horizon temporel de 25 ans soit long. Sauf si des données observées permettent aux auteurs de justifier les projections du modèle (5% de la cohorte toujours en vie à 15 ans), **il serait recommandé de réduire l'horizon temporel à 15 ans dans l'analyse de référence afin de réduire l'incertitude des extrapolations.**

Modélisation

- 12) Concernant la revue systématique de la littérature et la faisabilité de la méta-analyse (Annexe 6 du rapport technique), pourriez-vous mettre à jour l'annexe 4 avec les dernières publications d'ASPIRE et ENDEAVOR ?

- 13) Concernant la vérification de l'hypothèse des risques proportionnels.

- a. En plus de la figure rapportant les 4 courbes de Kaplan-Meier sur un même graphe (ex : pour l'OS, page 17/144 de l'annexe 7), pourriez-vous fournir un graphe pour chaque essai présentant les deux courbes de Kaplan-Meier avec les bandes de confiance et tester l'hypothèse sur les données individuelles des essais.

- b. Pourriez-vous commenter le graphique pour la SG entre KRd et RD (p.48 du rapport technique), le croisement des courbes, et discuter de l'impact d'une hypothèse des risques proportionnels critiquable sur les 6 premiers mois ?
 - c. Pourquoi ne pas avoir utilisé, dans l'analyse de référence, des analyses de Cox par morceaux pour tous les traitements (cf. Q13.b) ?
- 14) Concernant l'analyse groupée d'ASPIRE et ENDEAVOR (Annexe 7 du rapport technique)
- a. Pour quelles raisons les interactions de l'effet traitement avec « age » et « prior ASCT » n'ont-elles pas été incluses dans l'analyse finale du modèle de cox ?
 - b. Pourriez-vous présenter les distributions des caractéristiques des patients et faire un test global (en rapportant les valeurs de p) de la différence des distributions pour chaque caractéristique des patients présentée dans le tableau 3 p.10, et ce à minima pour l'âge, les greffes de cellules souches hématopoïétiques et le type de traitement reçu précédemment ?
 - c. Pourriez-vous nous fournir l'Appendix C (c.-à-d. le plan d'analyse statistique) de l'Annexe 7 dans un fichier séparé ?
- 15) Concernant le choix des fonctions paramétriques dans l'évaluation médico-économique :
- a. Pourquoi ne pas avoir modélisé l'extrapolation des résultats cliniques en utilisant une fonction paramétrique sur Vd (Log-logistique pour OS, Weibull pour la PFS) et des HR sur les autres traitements, tel que recommandé par le prestataire de l'analyse statistique (en Annexe 7) ?
 - b. Le choix de la fonction paramétrique pour extrapoler l'OS de Rd semble surestimer les résultats (Tableau 39 du rapport technique). Puisque la fonction Weibull présente les valeurs AIC/BIC les plus faibles (Tableau 13), et des prédictions réalistes à 5 et 8 ans (Tableau 14), il semblerait conservateur de l'utiliser en analyse de référence.
- 16) Que signifie la mention « population générale » dans les tableaux 13 et 14 ?
- 17) Que voulez-vous dire par « ajustement conjoint » sur les données Rd et KRd dans l'essai ASPIRE (p.43 du rapport technique) ?
- 18) Concernant les effets indésirables, la modélisation repose sur l'hypothèse implicite d'une incidence constante au cours de la prise de traitement.
- a. En quoi cette hypothèse est-elle recevable ?
 - b. La différence de fréquence des neutropénies entre Rd et Vd, découlant de cette hypothèse, est-elle réaliste (23,39% et 1,32%) ?
- 19) Les effets indésirables pris en compte sont définis sur la base d'une différence de fréquence d'occurrence entre KRd et Rd de 1% (page 52) et 2% (page 68). Quel taux avez-vous appliqué et sur quels effets indésirables cela a-t-il un impact ? Discuter de l'impact final de ce choix sur le RDCR.

Estimation des utilités

- 20) L'analyse de référence du modèle fait l'hypothèse que l'impact sur la qualité de vie des effets indésirables et des hospitalisations est capturé par les scores d'utilités dérivés du questionnaire QLQ-C30. Aucun décrétement supplémentaire n'est donc appliqué. Discuter du risque de sous-estimation de ces impacts en appliquant cette méthode.
- Explication de la question : si les questionnaires sont présentés à distance des événements considérés, ils peuvent ne pas les prendre en compte.*
- 21) Merci de développer les raisons pour lesquelles les données d'utilité issues de l'essai MM003 n'ont pas été considérées dans le modèle.
- 22) Comment expliquez-vous que le score de qualité de vie augmente après le premier cycle de traitement dans les deux bras (tableau 23, p.58) ?

Estimation des coûts

23) Il est attendu une correction dans l'analyse de référence en supprimant l'hypothèse de partage des flacons.

Explication de la question : Le RCP de Kyprolis® précise que « le délai entre la reconstitution et l'administration ne doit pas dépasser 24 heures. [...] Toutefois du point de vue microbiologique, le produit doit être utilisé immédiatement. ». L'hypothèse de partage de flacon ne semble donc pas réaliste au vu du RCP et de la taille de la population cible.

24) Concernant les coûts de suivi, pouvez-vous confirmer que vous avez appliqué les coûts estimés à partir des données de consommation recueillies au cours des périodes « start to end of treatment » de l'étude de consommation de soins (annexe 5) ? Pouvez-vous commenter ce choix ?

25) Pouvez-vous confirmer qu'aucun coût de suivi et d'hospitalisation n'est pris en compte en post-progression ? Détailler la méthode de calcul du coût de prise en charge post-progression.

Il est attendu une justification, au regard des pratiques actuelles, d'un coût de prise en charge post-progression inférieur au coût de prise en charge avant progression. En l'absence de cette argumentation, il est attendu qu'une correction des coûts de prise en charge soit apportée au modèle afin d'intégrer des coûts plus en rapport avec la réalité de la prise en charge.

Explication de la question : Le modèle tel que déposé dans sa première version utilise un coût par cycle en PFS bien plus élevé qu'en PPS ce qui semble contre-intuitif et en opposition avec les données présentées dans le tableau 43 p.60 de l'annexe 5. D'autre part, les évaluations du NICE dans le myélome multiple montrent aussi des consommations de ressources plus importantes en PPS qu'en PFS (ex : Table 22 p.68 rapport ERG TA171 ; Table 5.9 p.63 rapport ERG TA338). L'absence de coûts d'hospitalisation et de coûts de suivi dans l'état PPS du modèle ne semble pas réaliste.

26) Est-il possible de partager avec nous le fichier excel / SAS ayant permis de calculer les résultats de l'« étude consommation de soins 2014 » utilisés dans le modèle ?

27) Pourquoi les coûts des EI de grade 3 et 4 reproduits page 69 ne correspondent pas aux coûts présentés par EI dans l'étude « healthcare resource use » page 64.

28) Pourriez-vous reporter les codes GHM (« DRG, inpatient hospitalizations ») utilisés pour renseigner les coûts p.50 de l'annexe 5 ?

29) L'annexe 2 précise qu'en France 67% seulement des patients qui ont reçu une deuxième ligne de traitement reçoivent une troisième ligne, et 56% d'entre eux, une quatrième ligne. Ces proportions sont-elles reprises dans la valorisation des coûts de l'état de santé post-progression ?

30) Commentez la représentativité des patients inclus dans l'étude ayant servi à documenter les coûts dans le modèle par rapport à la population simulée ?

Résultats et analyses de sensibilité

31) Les RDCR doivent être calculés suivant la frontière d'efficience : Rd vs VD et Krd vs Rd

32) Il est attendu que les analyses de sensibilité probabilistes soient réalisées toutes options confondues et non deux à deux (méthode d'analyse de sensibilité multi-options, sur la base du bénéfice net).

Analyses de sensibilité supplémentaires

- 33) Il est attendu une analyse de sensibilité incluant les caractéristiques des patients de la population française atteinte de myélome multiple notamment en termes d'âge (ex : le tableau 20 de l'annexe 1 reporte un âge médian pour la population initiant la première ligne de traitement de plus de 70 ans).
- 34) Dans le modèle, la cohorte a une probabilité nulle de recevoir de la carfilzomib après 18 cycles. Or, d'après le RCP, « Le traitement par Kyprolis® peut être poursuivi jusqu'à la progression de la maladie ». Pourriez-vous intégrer en analyse de sensibilité une probabilité non nulle de recevoir le traitement entre le 18^{ème} cycle et la progression.
- 35) Concernant l'analyse de sensibilité proposant une calibration des données de survie de Rd sur les données du registre des hémopathies, présenter conjointement les durées OS et PFS selon les deux méthodes afin de pouvoir les comparer (seules les données issues du registre sont présentée p.49).
- 36) Il est attendu une analyse de sensibilité permettant de mesurer l'impact d'une décroissance importante de la qualité de vie dans l'état post-progression en testant une modification du différentiel 0.050 (par exemple en prenant le différentiel de 0.17 observé dans Van Agthoven).

Discussion

- 37) Pourriez-vous comparer et discuter les caractéristiques des patients inclus dans l'analyse de l'EGB (annexe 1), la chart review pour la France (annexe 5), avec celles des études ASPIRE et ENDEAVOR, notamment en termes d'âge (facteur pronostique), de traitements précédemment reçus, de transplantation de cellules souches hématopoïétiques. **Quelle conclusion en tirez-vous en termes de représentativité de la population simulée par rapport à la population traitée** en France et en termes d'impact sur les résultats de l'évaluation médico-économique ?
- 38) Pourriez-vous discuter des forces et limites de la méta-analyse telle qu'elle aurait été menée versus l'analyse groupée réalisée, en considérant que :
- a. La non-faisabilité de la méta-analyse entraîne l'exclusion d'un certain nombre de traitements recommandés par la HAS, l'INCa et l'ESMO pour les patients préalablement traités pour un myélome multiple tel que bendamustine, pomalidomide, thalidomide, les agents alkylants (melphalan, cyclophosphamide), empêchant ainsi de démontrer l'efficacité de Kyprolis® dans l'ensemble de la stratégie actuellement recommandée en deuxième et troisième ligne.
 - b. La fragilité de l'analyse groupée due à la différence considérable de temps de suivi entre les essais ENDEAVOR (12 mois) et ASPIRE (32 mois) et l'immaturité des données de survie de ENDEAVOR.

Validation

- 39) Pourriez-vous comparer et discuter les coûts moyens par patient obtenu dans l'étude de consommation de soins (annexe 5, p.66 + data consideration p.69) et les sorties du modèle ?
- 40) Pouvez-vous apporter une analyse critique des 3 études utilisées pour valider le choix de la fonction, en expliquant les différences importantes observées à 5 ans.
- 41) Le tableau S3 des annexes de la publication de l'étude ASPIRE (Stewart et al. 2015) reportent les causes d'arrêt de traitement dont la progression du myélome pour 40 % (bras traitement) et 50 % (bras contrôle) des patients. Pourriez-vous comparer ces résultats avec les sorties du modèle ?
- 42) L'annexe 5 « *Healthcare resource use of symptomatic multiple myeloma patients* » montre p.24 que les durées moyennes de traitement en France sont de 12 mois pour

lenalidomide et 8 mois pour bortezomib. Pourriez-vous comparer et discuter ces résultats avec les sorties du modèle ?

Bibliographie

Borget I, Antoun S, Chachaty E et al. Modalités de prise en charge des neutropénies fébriles chez les patients traités pour une néoplasie solide dans le service d'urgence de cancérologie Gustave-Roussy et leurs impacts économiques. *Bulletin du Cancer* 2014; 101: 925-31.

Dimopoulos et al "Carfilzomib and dexamethasone (Kd) vs bortezomib and dexamethasone (Vd) in patients (pts) with relapsed multiple myeloma (RMM): Results from the phase III study ENDEAVOR" *J Clin Oncol* 33, 2015 (suppl; abstr 8509).

Longworth L, Yang Y, Young T et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess* 2014; 18: 1-224.

Monnereau A, Remontet L, Maynadié M, Binder-Foucard F, Belot A, Troussard X, et al. Estimation nationale de l'incidence des cancers en France entre 1980 et 2012. Étude à partir des registres des cancers du réseau Francim Partie 2 - Hémopathies malignes. Saint-Maurice : Institut de veille sanitaire; 2013.

Palumbo A, Anderson K. Multiple myeloma. *The New England journal of medicine* 2011; 364: 1046-60.

Proskorovsky I, Lewis P, Williams CD et al. Mapping EORTC QLQ-C30 and QLQ-MY20 to EQ-5D in patients with multiple myeloma. *Health Qual Life Outcomes* 2014; 12: 35.

Ruggeri K. Estimating the Relative Effectiveness of Treatments in Relapsed/Refractory Multiple Myeloma through a Systematic Review and Network Meta-Analysis. December 3, 2015; *Blood*: 126 (23)

Signorovitch et al 2010. Comparative effectiveness without head-to-head trials. A method for matching-adjusted indirect comparisons applied to Psoriasis treatment with adalimumab or etanercept. Signorovitch JE et al. *Pharmacoeconomics* 2010.

Stewart AK, Rajkumar SV, Dimopoulos MA, Masszi T, Špička I. ASPIRE Investigators. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. *N Engl J Med*. 2015 Jan 8;372(2):142-52

Turini M, Piovesana V, Ruffo P, Ripellino C, Cataldo N. An assessment of chemotherapy-induced nausea and vomiting direct costs in three EU countries. *Drugs Context* 2015; 4: 212285.

Versteegh MM, Leunis A, Luime JJ, Boggild M, Uyl-de Groot CA, Stolk EA. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. *Med Decis Making* 2012; 32: 554-68.



Toutes les publications de la HAS sont téléchargeables sur
www.has-sante.fr

N° ISBN : 978-2-11-152287-9