

BILAN DES AVIS D'EFFICIENCE RENDUS AU 31 JUILLET 2015 – FOCUS SUR LES ASPECTS METHODOLOGIQUES

Document de travail

Fabienne MIDY¹, Véronique RAIMOND^{1,2}, Cléa SAMBUC¹, Catherine Rumeau-Pichon¹

✦ Personnes ayant collaboré aux avis

Membres experts de la CEESP : Alran Séverine, Audibert Martine ; Benamouzig Daniel ; Bideau Daniel ; Boelle Pierre-Yves ; Bonnel Galadriel ; Briot Martin ; Chauvin Pauline ; Clément Valérie ; Dervaux Benoît ; Duguet Christophe ; Dupont Jean-Claude K ; Favier Christian ; Frenkiel Jérôme ; Gerbaud Laurent ; Harousseau Jean-Luc ; Josselin Jean-Michel ; Lacoste Olivier ; Le Gales Catherine ; Lebrun Louis ; Lehr-Drylewicz Anne-Marie ; Martineaux Claude ; Milcent Carine ; Perrier Lionel ; Plan Pascal ; Remy Pierre-Louis ; Roudot-Thoraval Françoise ; Rusch Emmanuel ; Sannie Thomas ; Sermet Catherine ; Séror Valérie ; Tregoures Dominique ; Wittwer Jérôme.

Membres du SEESP : Affinito Sylvain ; Bongiovani Isabelle ; Couillerot Anne-Line ; Ghabri Salah ; Hamers Françoise ; Hirtzlin Isabelle ; Midy Fabienne ; Pibouleau Leslie ; Poullié Anne-Isabelle ; Raimond Véronique ; Sambuc Cléa ; Thébaut Clémence ; Zanetti Laura.

✦ Informations sur les auteurs

¹. Haute Autorité de Santé. 2 av du stade de France, 93 218 Saint Denis la Plaine

². Université Paris 1 Sorbonne, Hospinomics

Table des matières

1. INTRODUCTION : OBJECTIF ET METHODE	2
2. PROCESSUS D'ELABORATION DES AVIS D'EFFICIENCE	3
3. BILAN DES 30 PREMIERS AVIS VALIDES PAR LA CEESP AU 31 JUILLET 2015	5
4. ANALYSE METHODOLOGIQUE DES EVALUATIONS SOUMISES A LA CEESP	8
4.1. Objectif de l'évaluation	8
4.2. Choix structurants	9
4.3. Modélisation.....	11
4.4. Estimation des coûts et des utilités.....	15
4.5. Résultats et analyses de sensibilité	17
5. DISCUSSION	19
6. Conclusion	21

1. INTRODUCTION : OBJECTIF ET METHODE

Depuis le 3 octobre 2013, la Commission d'Évaluation économique et de Santé publique (CEESP) de la Haute Autorité de santé (HAS) émet des avis sur l'efficacité des produits de santé pour lesquels une amélioration du service médical rendu (ASMR I à III) ou du service attendu (ASA I à III) est revendiquée et pour lesquels il est attendu un « impact significatif sur les dépenses d'assurance maladie compte-tenu de [leur] incidence sur l'organisation des soins, les pratiques professionnelles ou les conditions de prise en charge des malades et, le cas échéant, de [leur] prix » [1][2][3]. Ces avis d'efficacité sont transmis au Comité économique des produits de santé (CEPS) qui détermine le prix du produit de santé et, pour le médicament, à l'Union nationale des caisses d'assurance maladie (UNCAM) qui en fixe le taux de remboursement.

L'objectif des avis rendus par la CEESP est de transmettre au CEPS une information pertinente et scientifiquement fiable sur l'efficacité des produits de santé visés par l'Art. R.161-71-1 du CSS [1][3].

Pour rédiger son avis, la CEESP se fonde sur l'évaluation coût-résultat du produit, réalisée par l'industriel, et sur l'expertise méthodologique qui en est faite par le Service d'Évaluation économique et de Santé publique (SEESP) de la HAS¹, en coordination avec des rapporteurs membres de la CEESP. L'expertise, exhaustive et détaillée, de l'évaluation soumise par l'industriel a pour objectif de vérifier la conformité méthodologique de l'évaluation, en particulier au regard des principes méthodologiques retenus par la CEESP [4].

En effet, la qualité méthodologique de l'évaluation réalisée détermine le degré de confiance attribuable aux résultats quantitatifs présentés. Il apparaît que de nombreux avis présentent des réserves méthodologiques importantes, associées à une forte incertitude sur les résultats. Cette incertitude limite considérablement la pertinence de l'information transmise par l'industriel et donc l'usage qui peut en être faite par le décideur [5][6].

Pour améliorer la qualité de l'information transmise au CEPS, il est nécessaire de réduire l'incertitude sur l'évaluation et donc d'augmenter la qualité méthodologique des dossiers déposés. Le SEESP a mis en place un suivi des expertises réalisées de manière à identifier et discuter les éléments de méthode qui sont le plus fréquemment observés par la HAS comme des limites de l'évaluation proposée et qui perdurent malgré l'échange technique entre l'industriel et la HAS.

Il est attendu que ce travail de recensement tende à une meilleure qualité des dossiers, à travers deux mécanismes :

- Permettre aux industriels d'anticiper plus précisément les prérequis méthodologiques et leur permettre de déposer d'emblée des dossiers plus conformes ;
- Permettre à la HAS d'intégrer des prescriptions méthodologiques dans la mise à jour du guide méthodologique [4] attendue fin 2016.

Au-delà de cette amélioration attendue, certaines limites ne peuvent être supprimées à court terme ou sont intrinsèques à l'exercice réalisé, et comprendre leur impact permettra une meilleure interprétation du résultat de l'évaluation, ce qui concourra à la montée en charge du calcul économique dans la décision en santé.

¹ Dans ce texte, le terme « évaluation » se réfère à l'analyse déposée par l'industriel auprès de la CEESP, et le terme « expertise » renvoie à l'analyse critique qui en est faite par le SEESP et la CEESP.

Objectif

Après un bilan synthétique des avis rendus par la CEESP sur l'efficacité des produits évalués, l'objectif de ce travail est de recenser et d'analyser les limites méthodologiques identifiées par la HAS dans les évaluations soumises par l'industriel, afin de développer des pistes d'amélioration.

Démarche méthodologique adoptée : source et outils

Au-delà d'une analyse rétrospective de l'activité réglementaire de la HAS, par ailleurs partiellement publiée dans le cadre de son rapport d'activité (HAS, 2015), il s'agit d'analyser de manière transversale les principales limites méthodologiques des travaux présentés par les industriels.

Entre le 1^{er} octobre 2013 et le 31 juillet 2015, la HAS a rendu 30 avis d'efficacité pour des produits de santé (5 déposés en 2013, 21 en 2014 et 4 en 2015).

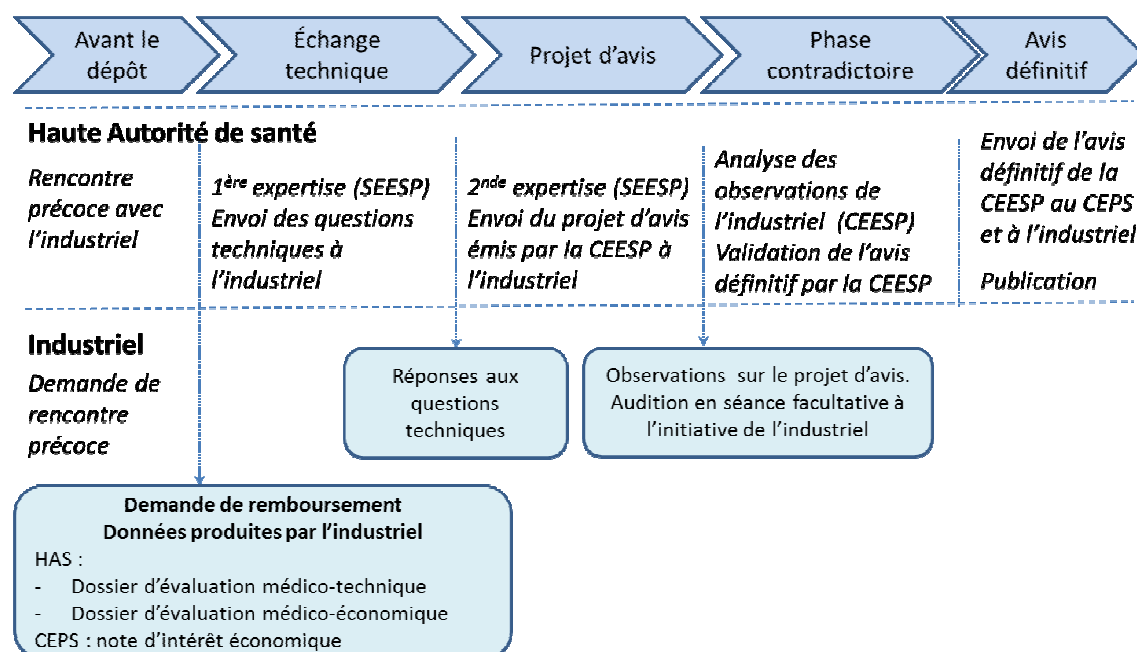
Le travail proposé repose sur un bilan des avis rendus, avec une analyse qualitative des conclusions émises par la CEESP sur l'efficacité des produits évalués, ainsi qu'une analyse transversale des questions méthodologiques traitées par l'industriel au cours de l'examen du dossier soumis, d'une part, et des réserves méthodologiques finalement émises d'autre part.

2. PROCESSUS D'ELABORATION DES AVIS D'EFFICACITE

► Description du processus d'élaboration d'un avis d'efficacité

Le processus mis en place par la HAS pour répondre à sa nouvelle mission (figure 1) repose sur trois principes ayant pour objectif la production d'une information utile à la fixation du prix par le CEPS : un échange constructif entre la HAS et l'industriel ; une expertise méthodologique exhaustive et détaillée par la HAS des données produites par l'industriel ; une analyse par la CEESP de ces données dans une perspective d'aide à la décision.

Figure 1 : Processus de production des avis d'efficacité



En amont de l'analyse par la CEESP des données produites, la HAS a prévu :

- une expertise méthodologique de l'évaluation économique déposée par l'industriel par le SEESP et la sous-commission économie, à laquelle la CEESP a délégué la validation de l'expertise méthodologique et technique des modèles produits par l'industriel ;
- deux échanges qui peuvent éventuellement être organisés avec les industriels : une rencontre précoce, à l'initiative de l'industriel, avant le dépôt du dossier et un échange technique au cours de l'instruction, à l'initiative de la HAS. Au cours de cet échange, le SEESP adresse une liste de questions techniques à l'industriel, lequel dispose alors d'un délai de 15 jours pour y répondre par écrit. Il peut également demander à présenter oralement ses réponses à la sous-commission économie. Cette phase d'échange est l'occasion pour les industriels de clarifier les choix méthodologiques qu'ils ont effectués, d'apporter des justifications sur des limites constatées par le SEESP et de fournir des analyses complémentaires le cas échéant.

Sur la base de ces éléments, un projet d'avis d'efficience est discuté et validé par la CEESP, puis transmis à l'industriel. Celui-ci dispose de 8 jours à compter de la réception du projet pour faire des observations écrites ou éventuellement faire une demande d'audition auprès du président de la CEESP. L'audition ne peut concerner que les éléments transmis par l'industriel au moment du dépôt de dossier ou de l'échange technique, aucune nouvelle donnée ou nouvelle analyse n'est prise en considération à ce stade de l'instruction d'un dossier.

A la suite de cette phase contradictoire, l'avis définitif de la CEESP est transmis à l'industriel et au CEPS. Il est ensuite publié sur le site internet de la Haute Autorité de santé dans un délai variable.

► Description d'un avis d'efficience

L'avis d'efficience est construit en trois parties, présentées ci-après [7]. Il est complété par quatre annexes présentant : le contexte de la demande déposée par l'industriel, l'expertise méthodologique par le SEESP des données transmises par l'industriel, un tableau de synthèse des réserves émises par la CEESP, et la liste de questions transmises à l'industriel lors de l'échange technique, le cas échéant.

Dans la première partie de l'avis, la CEESP évalue la conformité méthodologique de l'analyse coût-résultat proposée par l'industriel par rapport aux recommandations adoptées par la HAS en matière d'évaluation économique (Haute Autorité de santé, 2011).

Les réserves identifiées lors de l'expertise méthodologique de l'évaluation sont hiérarchisées selon trois niveaux :

- Réserve mineure (R⁻) : élément non conforme aux recommandations en vigueur, mais dont l'impact attendu sur les conclusions est jugé négligeable ou limité.
- Réserve importante (R⁺) : élément non conforme aux recommandations en vigueur, qui peut être justifié, mais dont l'impact attendu sur les conclusions est important, en particulier en termes d'incertitude.
- Réserve majeure (R⁺⁺) : élément non conforme aux recommandations en vigueur, qui invalide tout (réserve majeure globale) ou partie (réserve majeure partielle) de l'étude médico-économique.

La seconde partie de l'avis présente la conclusion de la CEESP sur l'efficience du produit, considérant le résultat de l'évaluation médico-économique présentée par l'industriel et le degré de confiance que l'on peut accorder à ce résultat. En cas de réserve majeure globale, la CEESP invalide l'analyse coût-résultat, ne présente pas les résultats estimés et conclut que l'efficience du produit n'est pas démontrée. Lorsque la réserve majeure est partielle, la CEESP invalide l'utilisation qui est faite du modèle ou une partie des résultats, mais n'invalide pas le modèle en lui-même. La CEESP est alors en mesure de produire un avis sur l'efficience du produit à partir, soit de résultats extraits de l'évaluation présentée par l'industriel, soit d'une analyse *de novo* réalisée par la CEESP. Dans tous les cas,

lorsqu'aucune réserve majeure globale n'est émise sur l'évaluation, la CEESP rapporte les résultats quantitatifs de l'évaluation. Elle explicite ensuite tous les éléments dont elle dispose pour interpréter ce résultat, en particulier en termes d'incertitude. Enfin, la CEESP peut interpeler le CEPS sur certains éléments de contexte lorsqu'elle le juge nécessaire.

Dans une troisième partie, la CEESP précise, le cas échéant, la nature des données (scientifiques) qui devront être produites, pour permettre une mesure plus robuste de l'efficacité lors de la réévaluation des produits.

3. BILAN DES 30 PREMIERS AVIS VALIDES PAR LA CEESP AU 31 JUILLET 2015

Au 31 juillet 2015, 37 dossiers économiques ont été déposés, dont la majorité concernent un médicament (3 concernent un dispositif médical). Trente avis ont été définitivement validés et transmis au CEPS. Les autres dossiers étaient en cours d'instruction au moment de la rédaction de cette analyse.

La majorité des dossiers concernent des demandes de primo-inscription (23/30). Les demandes de réévaluation ou réinscription concernent 4 dossiers et les demandes d'extension d'indication concernent également 4 dossiers².

► Échanges avec les industriels

Une rencontre précoce a été organisée à la demande de l'industriel pour 19 dossiers. Cependant, dans cette phase de mise en place, les rencontres n'ont permis qu'une intégration à la marge de préconisations méthodologiques (14/19 rencontres se sont déroulées moins de 4 mois avant le dépôt de dossier).

Pour que cette rencontre constitue un apport réel pour l'industriel, il faudrait que celui-ci demande à rencontrer le SEESP très en amont, dès la fin des études de phase II, afin d'optimiser les données recueillies lors des essais de phase III (choix du comparateur, recueil de données d'utilité, etc.).

En cours d'instruction, la totalité des dossiers a fait l'objet d'un échange technique sous la forme d'une liste de questions rédigées par le SEESP et auxquelles l'industriel a apporté des réponses écrites. Pour 12 dossiers, l'industriel a souhaité présenter ses réponses devant la sous-commission économie.

Le volume de questions transmises à l'industriel est très variable, mais la tendance a été une nette augmentation de leur nombre sur la période. Principalement, les questions visent à apporter une justification à un choix méthodologique, à expliquer un calcul ou une démarche, ou à refaire des calculs (modification de l'analyse de référence ou analyses de sensibilité supplémentaires).

Enfin, après la rédaction du projet d'avis, 27 avis ont fait l'objet d'observations écrites par l'industriels, parmi lesquels 14 ont également sollicité une audition auprès de la CEESP.

► Conformité méthodologique des évaluations soumises à la CEESP

Sur les 30 dossiers analysés, 10 avis ont été rendus avec une réserve majeure globale invalidant l'ensemble de l'étude (cf. Tableau 1).

Cela signifie que, dans la majorité des dossiers (20/30), la CEESP a pu documenter le ratio différentiel coût-résultat (RDCR) du produit évalué par rapport à un ou plusieurs

² Un dossier revendiquait une demande de réévaluation et une demande d'extension d'indication.

comparateurs. Le RDCR permet de mesurer le surcoût associé au gain de santé généré par un produit plus coûteux et plus efficace que ses comparateurs. Le résultat est généralement exprimé en coût par QALY gagné, le QALY désignant « l'année de vie pondérée par la qualité de vie ».

Tableau 1 : Bilan des réserves émises par la CEESP

Nature de la réserve	Nombre
Réserve majeure globale	10/30
Réserve majeure partielle	4/30
Réserve importante	15/30
Réserve mineure	1/30
Aucune réserve	0

► Conclusions de la CEESP sur l'efficacité des produits

Dans le cas d'une réserve majeure, lorsque cela était possible, la CEESP a tenté d'apporter au CEPS une information pouvant aider à la fixation des prix, malgré l'absence de résultat chiffré présenté. Par exemple, elle a souligné, pour un dossier, le caractère extrêmement élevé du RDCR estimé par l'industriel, s'il devait être confirmé par une étude de bonne qualité. Pour un autre dossier, elle a indiqué que les différences de coûts de traitement entre les deux stratégies étaient interprétables.

Dans le cas particulier d'une réserve majeure partielle, le modèle n'étant pas invalidé, la CEESP peut redéfinir l'analyse de référence en adoptant des choix méthodologiques en cohérence avec son guide méthodologique (définition d'une population d'analyse adaptée à l'indication ; définition de stratégies de traitement adaptées à la pratique courante) afin de produire des résultats quantitatifs pertinents.

Sur les 20 avis sans réserve majeure globale, les RDCR estimés par les industriels sont très variables. Deux produits étaient présentés comme dominants (plus efficaces et moins coûteux que le comparateur). Les ratios différentiels coûts-résultats des autres produits étaient très hétérogènes, variant de moins de 10 000€ par QALY à plus de 200 000€ par QALY, avec une moitié de produits pour lesquels le RDCR était supérieur à 50 000€ par QALY (11/20). Sont majoritairement concernés par des RDCR élevés : les produits en oncologie (3/3) et les produits traitant l'hépatite C (3/4). Les évaluations portant sur des vaccins, *a contrario*, ont présenté des RDCR inférieurs à 30 000 €/QALY³.

Compte-tenu des limites méthodologiques associées à ces résultats, l'un des objectifs fondamentaux de la conclusion de la CEESP est d'en proposer une interprétation, qui précise le degré d'incertitude qui les entoure. Ainsi, la CEESP a indiqué une probable sous-estimation du RDCR pour 10 dossiers, mentionné un niveau de confiance faible avec un risque d'observer un RDCR d'une valeur très différente en pratique courante pour 3 dossiers et souligné l'intérêt limité du résultat pour 4 dossiers (analyse non adaptée au problème de décision, population de l'indication non intégralement couverte, comparateur peu utilisé en pratique). Pour 6 dossiers, la CEESP disposait de suffisamment d'éléments pour réaliser une nouvelle évaluation du produit : pour 3/6 dossiers, cette évaluation a été retenue en analyse de référence et seuls ses résultats ont été présentés dans l'avis ; pour les autres 3/6 dossiers, la CEESP a présenté des résultats alternatifs, sans pour autant aller jusqu'à rejeter l'analyse retenue par l'industriel.

Un élément essentiel de l'interprétation du RDCR est le prix du produit. L'analyse de référence est fondée sur le prix revendiqué par l'industriel et aucune analyse *a posteriori* de

³ Il est à noter que deux des vaccins concernés ont reçu une SMRi de la Commission de la Transparence.

l'efficacité des produits n'est réalisée, lorsque le prix facial négocié est différent du prix revendiqué. Afin de pouvoir éclairer le CEPS sur l'impact d'une variation de prix sur l'efficacité, la CEESP requiert que différents prix soient simulés dans l'étude fournie par l'industriel. La conclusion de la CEESP indique systématiquement la variation du RDCR associée à une variation du prix du produit. Le prix du comparateur est également un élément fondamental de construction du RDCR, et sa variation, particulièrement probable lorsque le brevet arrive à échéance, peut fortement modifier le résultat, ce qui a été souligné pour 2 dossiers.

Enfin, la CEESP a souligné dans certaines de ses conclusions des éléments qu'elle juge importants à considérer dans la négociation des prix : l'existence d'un débat de société autour du produit évalué ou de la pathologie considérée, l'hétérogénéité des résultats en fonction des sous-populations, l'arrivée prochaine sur le marché de comparateurs cliniquement pertinents ou l'importance de considérer dans les évaluations les stratégies non médicamenteuses, l'impact économique de certaines stratégies industrielles (revendication d'un prix unique, quels que soient la durée de traitement ou le dosage ; retrait de la version antérieure du produit). Elle souligne également régulièrement l'importance des pratiques (prescription, organisation des soins, courbe d'apprentissage, etc.) qui peuvent limiter la transposition des résultats en pratique courante.

4. ANALYSE METHODOLOGIQUE DES EVALUATIONS SOUMISES A LA CEESP

L'analyse des conclusions de la CEESP permet de constater l'importance de l'incertitude liée à des limites méthodologiques. Le tableau ci-dessous synthétise l'ensemble des réserves émises par la CEESP. L'analyse développée ci-après sur les éléments qui ont soulevé le plus de réserves, ou ceux qui constituent un enjeu particulier, vise –outre leur identification– à discuter de leur impact pour l'évaluation et à proposer certaines pistes d'amélioration.

Tableau 2 : Synthèse des réserves émises par élément de l'évaluation

	Pas de réserve	Réserves mineures	Réserves importantes	Réserves majeures
Objectif de l'évaluation	27	0	4	0
Choix structurants de l'évaluation				
Analyse économique	27	1	1	1
Population d'analyse	23	3	2	2 partielles 1 globale
Perspective	25	5	0	0
Comparateurs	9	10	15	5
Horizon temporel	24	3	3	0
Choix méthodologique et hypothèses de modélisation				
Type de modèle	29	1	0	0
Structure	22	10	8	0
Hypothèses simplificatrices	17	17	8	0
Dimension temporelle	22	4	6	0
Population simulée	21	6	2	1
Données source et estimation des probabilités	4	20	14	7
Estimation des utilités et des coûts				
Estimation des utilités	1	37	15	2
Estimation des coûts	3	57	11	0
Résultat et analyses de sensibilité				
Résultats	18	9	6	5
Analyses de sensibilité	4	28	27	2 partielles 3

Note : la catégorisation des réserves reprend la présentation de l'analyse critique retenue par la HAS.

4.1. Objectif de l'évaluation

Identification

La définition de l'objectif de l'évaluation a conduit la CEESP à émettre une réserve importante dans quatre dossiers (4/30). Pour certains produits, la CEESP a été amenée à préciser la question à laquelle l'évaluation permettait finalement de répondre, sans émettre de réserves, considérant les comparateurs retenus ou l'indication considérée (6/30).

Explication

La définition de l'objectif de l'évaluation est un élément fondamental de la conception de l'évaluation, dans la mesure où l'objectif détermine les choix méthodologiques qui seront

faits. Or, définir l'objectif de l'évaluation ne peut se réduire à l'objectif opérationnel fréquemment avancé par les industriels (« calculer le RDCR du produit »).

En effet, selon la nature du produit évalué, sa place dans la stratégie thérapeutique et le moment de son évaluation, la question de l'efficacité peut être posée différemment. Il s'agit en particulier de préciser la place attendue du produit dans une prise en charge sanitaire (p.ex. en adjonction ou en substitution par rapport à une pratique existante) et la nature de la revendication (p.ex. première inscription, réévaluation ou réinscription, extension d'indication ; augmentation ou maintien d'un niveau de prix).

Par ailleurs, l'adéquation de l'objectif au contexte de la demande et aux choix méthodologiques réalisés dans le cadre de l'évaluation est généralement peu discutée par l'industriel. Par exemple, considérant l'absence de plusieurs comparateurs, la CEESP « souligne que l'évaluation présentée ne permet pas de documenter l'efficacité du produit dans la stratégie thérapeutique. [En effet,] rien ne garantit que les produits introduits dans l'évaluation soient sur la frontière d'efficacité [si tous les produits pertinents sont pris en compte], ce qui constitue une limite à l'interprétation des résultats en termes d'efficacité ».

Les réserves importantes émises ont porté sur la présentation d'une évaluation construite comme si le produit revendiquait une primo-inscription, alors qu'il était déjà commercialisé (réévaluation ou extension d'indication).

Pistes d'amélioration

Lorsqu'un produit est déjà commercialisé, qu'il dispose d'une indication et d'un prix, il est nécessaire de définir deux objectifs. L'évaluation doit permettre : d'une part, de documenter l'efficacité dans le contexte de la revendication en cours versus ses comparateurs et, d'autre part, de documenter en termes d'efficacité une éventuelle évolution entre la primo-inscription et la réinscription/réévaluation ou l'extension⁴, afin de pouvoir éclairer une décision de révision ou de maintien du prix existant. Par exemple, il s'agira de savoir si l'efficacité attendue, qui a présidé à la fixation du prix en primo-inscription, est confirmée en réinscription par de nouvelles données. Lorsqu'une demande d'extension d'indication est formulée dans une aire thérapeutique, l'évaluation doit permettre d'en documenter l'efficacité en rapportant l'ancienne indication à la nouvelle (cf. note 3).

4.2. Choix structurants

Parmi les choix structurants de l'évaluation économique, deux choix méthodologiques soulèvent des difficultés particulières : la définition de la population d'analyse et le choix des comparateurs.

Tableau 3 : Détail des réserves majeures sur les choix structurants de l'évaluation économique

Analyse économique	Absence d'analyse en coût-utilité
Population d'analyse	Calcul d'un RDCR sur la population d'analyse globale (Réserve majeure partielle). Les RDCR sur les sous-populations sont recevables. Calcul d'un RDCR sur la population d'analyse globale (Réserve majeure). Extrapolation du résultat à une population élargie (Réserve majeure partielle). Le RDCR est recevable sur une sous-population.
Comparateurs	L'évaluation intègre des comparateurs qui ne s'adressent pas aux mêmes groupes de patients Absence de comparateurs pertinents (4 Réserves majeures)

⁴ Il s'agit ici d'une demande d'extension dans une même aire thérapeutique, pas dans une nouvelle aire thérapeutique qui serait alors considérée comme une primo-inscription.

► Définition de la population d'analyse

Identification

La population d'analyse correspond à la population sur laquelle l'évaluation de l'efficacité est revendiquée. L'analyse des questions techniques met en évidence trois cas de figure de définition inadaptée de la population d'analyse.

1. La définition d'une population d'analyse globale est adaptée, mais elle doit être complétée par des analyses en sous-populations (3 dossiers).
2. La définition de la population d'analyse est adaptée, mais l'analyse économique requiert la prise en compte des seules sous-populations, avec la production de RDCR spécifiques à chacune. Dans certaines situations, l'estimation d'un RDCR sur la population d'analyse globale est non recevable (4 dossiers en échange technique).
3. Le périmètre de la population d'analyse est inadapté considérant la demande revendiquée (3 dossiers).

Explication

Les deux premiers cas renvoient à la problématique de l'hétérogénéité dans les évaluations économiques. Dans le premier cas, elle peut être traitée au sein d'une population globale avec des analyses en sous-population (p.ex. : efficacité moyenne estimée au sein d'une population hétérogène). L'hétérogénéité est limitée aux paramètres introduits dans le modèle, et ne remet pas en cause l'homogénéité des choix structurants de l'évaluation entre les différentes sous-populations. En l'absence de cette homogénéité, l'estimation d'un RDCR global n'a pas de sens (p.ex. : comparateurs différents selon les sous-populations). En conséquence, l'absence de sous-population d'analyse n'a pas la même portée : dans le premier cas, l'information produite est incomplète ; dans le second cas, l'information produite n'est pas recevable.

Concernant le périmètre de la population d'analyse, il devrait dans l'idéal correspondre à la population cible définie par l'industriel dans sa demande de remboursement. En pratique, elle peut différer de l'indication définie dans l'autorisation de mise sur le marché (AMM). Par exemple, lorsque l'indication d'AMM s'appuie sur une extrapolation, par l'EMA, des résultats cliniques observés dans les essais à une population plus large que celle identifiée initialement par l'industriel. L'industriel peut alors choisir de déposer un dossier sur la population cible définie dans l'AMM, mais il ne dispose alors pas de données suffisantes pour produire une évaluation économique sur cette partie de l'indication. L'industriel peut également choisir de restreindre sa revendication d'inscription au remboursement. Les limites associées à la définition du périmètre de la population d'analyse ont été, en général, corrigées ou justifiées au cours de l'échange technique.

Pistes d'amélioration

Les limites identifiées concernant la définition de la population d'analyse renvoient à une meilleure compréhension de l'évaluation réalisée et à la nécessaire cohérence entre la définition des choix structurants de l'évaluation et l'objectif de l'évaluation. Dans tous les cas, l'efficacité ne peut être validée que pour la population qui a été analysée dans l'évaluation économique. Il est donc important que l'industriel évalue la taille de la population d'analyse par rapport à la population concernée par l'AMM.

► Le choix des comparateurs

Identification

Sur les 30 dossiers analysés, aucune réserve majeure n'a été émise sur les comparateurs inclus dans l'évaluation. Pourtant, le choix des comparateurs dans l'analyse économique est une des dimensions qui soulèvent le plus de questions techniques (28/30) et le plus de réserves (21 dossiers sur 30, dont 16 dossiers avec au moins une réserve importante).

L'analyse des comptes rendus des discussions en sous-commission Economie et en CEESP permet d'apporter un premier éclairage à cette apparente contradiction. Il est en effet fréquemment mentionné, qu'au-delà du principe méthodologique appelant l'identification de toutes les options médicalement pertinentes, des contraintes opérationnelles sont à prendre en compte concernant l'introduction de certains comparateurs dans l'évaluation (ex. produits utilisés hors AMM ou produits dont le développement clinique est concomitant à celui qui est évalué). L'échange technique permet le plus souvent d'apporter des éléments de justification complémentaire. Un second éclairage est apporté par l'analyse de l'impact des principales limites importantes identifiées par la CEESP, à savoir l'absence de comparateurs pertinents (7/30) et un défaut de définition des comparateurs retenus (3/30).

Explication

Dans les deux cas, la limite identifiée n'invalide pas l'évaluation économique, mais en réduit la portée. Ainsi, lorsqu'il manque au moins un comparateur cliniquement pertinent, l'évaluation ne permet pas de documenter l'efficacité du produit dans l'ensemble de la stratégie thérapeutique, puisque l'introduction d'un comparateur peut modifier de manière significative le RDCR du produit ou même sa place sur la frontière d'efficacité. Pour autant, le RDCR apporte une information utile en documentant le différentiel de coût et de résultat avec le comparateur retenu. Dans le cas d'un défaut de définition des comparateurs, cela génère une incertitude importante sur la transposition des résultats en pratique courante. Par exemple, faire référence à la notion très générale de soins de support sans préciser à quelles modalités de prise en charge ces soins renvoient, ne permet pas de vérifier que cela correspond à une pratique ayant cours en France.

Ces évaluations permettent donc de produire une information partielle, qui reste plus utile au décideur que l'absence d'information, en situant le produit par rapport à certaines des options envisageables en termes de coûts et de résultats de santé. Cette volonté de ne pas censurer des informations utiles, même imparfaites, explique l'émission de réserves importantes et non majeures.

Pistes d'amélioration

Il est important qu'une distinction soit faite entre l'identification indispensable de toutes les options pertinentes et l'intégration souhaitable de toutes ces options dans l'évaluation. En termes de faisabilité, il n'est pas toujours possible d'inclure toutes les options envisageables dans l'évaluation ; encore faut-il que cette impossibilité soit justifiée. Tout comme pour la définition de la population d'analyse, l'enjeu ici est d'interpréter correctement l'évaluation réalisée. Connaître l'ensemble des options possibles par rapport aux options finalement intégrées dans le modèle va permettre de caractériser l'efficacité évaluée : efficacité dans la stratégie thérapeutique, efficacité de l'adjonction d'un produit dans un bras de traitement, efficacité de la substitution du produit à un autre, etc. Cela constitue une information pertinente pour caractériser le risque pris dans la décision, en fonction de l'importance des comparateurs manquants.

4.3. Modélisation

Sans négliger l'importance qui doit être accordée à la structure du modèle, trois aspects méthodologiques de la modélisation seront plus précisément analysés, en raison des réserves majeures émises (la population simulée, les données cliniques et l'estimation des probabilités), ou de la portée du choix méthodologique sur l'évaluation (la gestion de la dimension temporelle).

► Structure de la modélisation

Identification

Le choix du type de modèle ne soulève quasiment pas de questions techniques, ni de réserves. Concernant la structure du modèle et les hypothèses de modélisation qui sont faites, les questions techniques sont principalement de l'ordre de la justification et de l'explicitation (définition des états du modèle ; modélisation des transitions ; modélisation des événements ; modélisation de l'algorithme de traitement).

Une analyse plus précise des questions techniques et des réserves émises permet d'identifier deux éléments dont la modélisation soulève davantage de difficultés : les effets indésirables et les algorithmes de prise en charge (séquences de traitement, durée de traitement).

Explication

Ce constat permet de faire un premier commentaire assez général sur la qualité des modèles réalisés et la difficulté de présenter une modélisation avec un degré de précision et de détail suffisant, pour en permettre une bonne compréhension et une expertise rigoureuse et éclairée. Les réponses apportées par l'industriel permettent dans la plupart des cas de clarifier les choix réalisés concernant la structure du modèle.

Les limites identifiées dans la modélisation des effets indésirables et des arrêts de traitement sont plus difficiles à résoudre *a posteriori*. Or, les événements indésirables peuvent avoir un impact important sur le RDCR par leurs conséquences sur les coûts et sur la qualité de vie. De même, la modélisation des arrêts de traitement soulève régulièrement des demandes d'explication sur la manière dont ils sont pris en compte dans le modèle et a conduit à des réserves importantes dans 4 dossiers.

Pistes d'amélioration

Un certain nombre de questions et de réserves mineures peuvent être résolues par une amélioration de la rédaction du rapport technique, en lien avec la transmission du modèle électronique. En revanche, la modélisation des effets indésirables et des algorithmes de traitement, en particulier pour la prise en compte des arrêts de traitement, mérite d'être améliorée.

► La population simulée

Identification

Dans la majorité des dossiers (20/30), la cohorte simulée correspond à la population incluse dans les essais pivots réalisés en vue de l'obtention de l'autorisation de mise sur le marché. Dans les dossiers concernant un vaccin, la population simulée est caractérisée comme la population cible de la vaccination, à partir des données INSEE (4/30). Dans les autres cas, les caractéristiques de la population simulée peuvent être issues d'une étude observationnelle ou de sources multiples.

La définition de la population simulée fait l'objet de nombreuses questions lors de l'échange technique, relatives à la description insuffisante de la population simulée (13/30) et à la représentativité de la population simulée par rapport à la population susceptible d'être traitée en pratique courante (9/30).

Explication

Ces questions reflètent une réelle problématique méthodologique. Dans l'idéal, la population simulée devrait correspondre aux caractéristiques de la population réelle traitée en France, ou susceptible d'être traitée, pour favoriser la transposition des estimations en pratique courante et informer le décideur sur des résultats aussi proches que possible du contexte de la décision. Cependant, l'attribution de caractéristiques moyennes observées en pratique

courante à la population simulée dans les modèles de cohorte⁵ introduit une distorsion avec les données d'efficacité utilisées, et génère une incertitude difficile à explorer. L'approche la plus couramment mise en place dans les dossiers, consistant à simuler la population de l'essai clinique, est acceptable considérant les données disponibles en primo-inscription.

Cette situation, associée à une forte proportion de dossiers déposés lors d'une première inscription au remboursement, explique que la représentativité de la population simulée par rapport à la population traitée ou susceptible d'être traitée est un point qui soulève fréquemment des réserves après l'échange technique (2 réserves importantes et 5 réserves mineures). Ces réserves reflètent l'incertitude générée par ces choix méthodologiques, le plus souvent non quantifiable, sur la transposition du RDCR estimé en pratique courante.

Pistes d'amélioration

La simulation d'une cohorte ayant les caractéristiques de la population incluse dans les essais doit absolument s'accompagner d'une analyse approfondie des différences avec les patients traités en France et d'une discussion de leur impact sur l'efficacité attendue en pratique courante.

► Les données cliniques et l'estimation des probabilités

L'évaluation de l'efficacité du produit par rapport à ses comparateurs est le socle de toute évaluation médico-économique.

Identification

La majorité des dossiers ont fait l'objet de questions techniques (23/30) sur les données cliniques entrées dans le modèle, tant sur les sources retenues pour documenter l'efficacité ou la tolérance du produit ou du comparateur, que sur les méthodes qui ont permis de les intégrer dans le modèle.

Le niveau de qualification des réserves émises traduit l'enjeu que représente la disponibilité de données cliniques robustes pour l'évaluation économique. Les réserves portent sur la qualité des données cliniques (4 réserves majeures et 8 réserves importantes) et sur les méthodes appliquées aux données brutes pour les adapter à la modélisation économique (3 réserves majeures et 3 réserves importantes).

Tableau 4 : Détail des réserves majeures sur les données d'efficacité

Niveau de preuve des données de l'essai clinique	Essai de phase III avec cohorte historique, dont le critère principal de jugement est évalué à 100 jours. Analyse en sous-groupe d'un essai randomisé réalisé en ouvert, sans contrôle du biais de sélection par un score de propension. Essai de doses versus placebo n'apportant pas de données comparatives versus traitement actif comparateur. Sous-population : faible effectif dans l'essai (9 cas, 14% de l'essai).
Méthodes	Paramétrage des courbes de survie fondé sur un comparateur cliniquement non pertinent, en supposant la même forme pour les autres comparateurs Utilisation d'une méthode de comparaison naïve Utilisation d'une méthode de comparaison naïve

Explication

Le nombre important de questions posées lors de l'échange technique traduit la vigilance du SEESP concernant la fiabilité des données d'efficacité brutes et des données transformées introduites dans le modèle. Elle traduit également un défaut de précision récurrent dans les dossiers déposés initialement, tant sur les essais cliniques (description de l'essai clinique,

⁵ Cette limite ne concerne pas les modèles de niveau individuel fondés sur des équations de risque.

méthodes statistiques utilisées, données manquantes, choix du comparateur) que sur les autres sources de données utilisées (méthode de recherche documentaire et de sélection des données, choix d'une source de données lorsque plusieurs existent). De même, les méthodes d'estimation des probabilités intégrées dans les modèles ou les méthodes de calibration des données sont souvent insuffisamment détaillées.

Si l'échange technique permet généralement de combler un défaut de justification ou de description des données entrées dans le modèle, les réserves émises traduisent dans certains cas l'impossibilité de corriger certaines limites des données sources. Cela génère une incertitude importante, voire majeure, associée en particulier au manque de robustesse des essais (effectifs faibles, données non comparatives, données en sous-groupe issues d'analyse post-hoc).

Pistes d'amélioration

Si les données pour le produit évalué sont en général issues des essais développés par l'industriel, les méthodes de recherche et de sélection des données utilisées pour les comparateurs peuvent être améliorées.

Par ailleurs, il est nécessaire que les données cliniques sur lesquelles repose l'estimation de l'efficacité et de la tolérance du produit soient suffisamment matures pour fonder une évaluation économique. Enfin, les méthodes appliquées aux données brutes pour les adapter à la modélisation économique pourraient être améliorées, en particulier par le développement des méta-analyses en réseau.

► La gestion de la dimension temporelle

Identification

La gestion de la dimension temporelle dans la modélisation renvoie à trois éléments méthodologiques : la durée des cycles, la durée de simulation et les méthodes d'extrapolation à long terme des données.

La durée des cycles ne suscite quasiment jamais de questions lors de l'échange technique ou de réserves. En revanche, la durée de simulation et l'extrapolation à long terme des données cliniques disponibles a suscité des questions dans presque la moitié des dossiers (14/30). Suite à l'échange technique, le nombre de réserves est finalement limité, tant pour la durée de simulation (4 réserves, dont 3 réserves importantes) que pour les méthodes d'extrapolation (5 réserves, dont 3 importantes).

Explication

Les choix méthodologiques concernant la dimension temporelle ont des conséquences importantes sur l'interprétation de l'efficacité. D'une part, la valeur du RDCR peut être considérablement modifiée en fonction de la durée sur laquelle sont estimés les coûts et les résultats de santé. D'autre part, l'extrapolation à long terme des effets du traitement observés dans les essais cliniques, généralement de courte durée, est une source importante d'incertitude.

L'analyse des échanges avec l'industriel et des réserves émises par la CEESP montre qu'il existe un réel défaut de justification des techniques statistiques d'extrapolation appliquées ou des hypothèses d'extrapolation retenues (p.ex. l'application sur près de 30 ans d'une hypothèse de maintien d'un effet traitement observé sur quelques semaines).

Afin de limiter le risque d'une simulation artificiellement trop longue de la durée de vie, le modèle définit en fait une durée de simulation maximale dans la moitié des évaluations portant sur un horizon temporel vie entière (9/18 modèles avec un horizon vie entière). Notons que, si une proportion significative de patients est encore en vie, ce choix méthodologique revient à remettre en cause le choix structurant de l'horizon temporel, puisque la simulation s'arrête avant que la totalité des individus simulés soient décédés. Les

réserves émises signalent que la durée de simulation maximale retenue par l'industriel est trop longue, compte tenu des connaissances sur l'espérance de vie des patients, ce qui a pour conséquence de sous-estimer le RDCR du produit.

Pistes d'amélioration

Cette dimension de l'évaluation implique un arbitrage entre une durée de simulation suffisamment longue pour prendre en compte toutes les conséquences du produit et l'incertitude générée par l'extrapolation. Expliciter cet arbitrage permettrait de fonder le choix de la durée de la simulation.

En attendant un recul suffisant pour disposer de données cliniques observées, il est fondamental que les choix fondant les extrapolations soient parfaitement explicités et justifiés. Compte tenu de l'incertitude générée, il est également indispensable qu'un travail de validation de la durée de simulation et de l'extrapolation soit apporté par les industriels à partir de données en vie réelle, lorsqu'elles existent. L'analyse a montré un réel apport de l'échange technique sur cette question, plusieurs industriels ayant fourni des simulations complémentaires en réponse aux questions posées par la HAS.

4.4. Estimation des coûts et des utilités

► **L'estimation des utilités**

Identification

Comme dans le cas des données d'efficacité, des lacunes sont fréquemment relevées dans les dossiers initialement déposés concernant les sources de données : source retenue non décrite, recherche documentaire non présentée, sélection des sources retenues non justifiée ; ce dernier point suscitant un nombre important de questions techniques (19/30). L'intégration des scores d'utilité dans le modèle doit également souvent être clarifiée lors de l'échange technique (16/30).

Pour 4 dossiers, l'échange technique demandait explicitement de modifier l'analyse de référence sur les valeurs d'utilité. L'objectif de la demande du SEESP était d'intégrer les désutilités associées aux effets indésirables ou aux hospitalisations, de modifier la source retenue pour privilégier des données françaises et de supprimer une hypothèse non validée en faveur du produit.

Explication

Les sources pour documenter les scores d'utilité introduits dans les évaluations sont principalement la littérature et les études cliniques réalisées sur le produit. Si l'échange technique permet le plus souvent de mieux documenter la phase d'identification et de sélection des données, certaines limites perdurent et les réserves émises indiquent que la valorisation des états du modèle par un score d'utilité est une source d'incertitude fréquente. En particulier, la multiplicité d'états à valoriser et de comparateurs nécessite souvent le recours à des sources de données hétérogènes, à des hypothèses ou à des avis d'experts. L'incertitude générée par le choix des sources est difficilement réductible en l'absence de données issues d'une source unique et fiable. L'absence de justification ou de détails sur la source des données retenues est le principal motif de réserves (10 réserves mineures). L'incertitude générée peut conduire à une remise en cause de la source ou de la valeur retenue, par exemple lorsque des résultats contre-intuitifs sont constatés, ou lorsque le choix est favorable au produit sans justification (4 réserves importantes).

Quelle que soit la source retenue (étude ad hoc ou littérature), l'expertise des scores d'utilité intégrés dans le modèle nécessite que le SEESP dispose d'une description détaillée des méthodes qui ont permis d'estimer les scores : méthode de recueil (p.ex. : description de l'échantillon de patients, estimation du taux de réponse et traitement des données

manquantes, questionnaire) et méthode de valorisation (principalement la matrice de valorisation si le questionnaire EQ-5D est utilisé, et la méthode de « mapping » dans les autres cas). Les réserves émises sur les méthodes d'élicitation des scores sont fonction de l'incertitude qui est générée. Par exemple, l'utilisation de la matrice de valorisation anglaise est jugée comme générant une incertitude faible (4 réserves mineures). A contrario, l'utilisation d'une méthode non valide est sanctionnée par une réserve majeure (2 réserves majeures).

De la même manière, les réserves émises concernant l'intégration des scores d'utilité dans le modèle dépendent de l'impact attendu sur le résultat de l'évaluation. Les réserves émises sont importantes si l'incertitude générée est potentiellement élevée (5 réserves importantes, en particulier sur la non prise en compte de l'impact sur la qualité de vie des événements indésirables).

Pistes d'amélioration

Considérant le manque de disponibilité des données, un arbitrage est souvent nécessaire entre exigence méthodologique et opérationnalité. Cependant, il est nécessaire que l'incertitude générée par cet arbitrage soit documentée, en justifiant le choix des sources de données retenues pour documenter les scores d'utilités, d'une part, et en réalisant des analyses de sensibilité permettant d'appréhender l'impact attendu du niveau d'utilité sur les résultats du modèle, d'autre part. Le nombre d'analyses de sensibilité complémentaires demandées au cours de l'échange technique montre que des efforts sont nécessaires pour documenter l'incertitude générée par les scores d'utilité.

► **L'estimation des coûts**

Identification

Le nombre de questions sur les coûts posées lors de l'échange techniques est très important, car les demandes de justification ou de précision sont quasiment systématiques. Suite à l'échange technique, le nombre de réserves reste très important (68 réserves pour 30 dossiers), mais elles sont principalement mineures (84% des réserves sont mineures).

L'échange technique est également souvent l'occasion de signaler des erreurs ou incohérences dans l'estimation des coûts.

Explication

Les questions techniques et les réserves émises portent sur tous les domaines de l'évaluation des coûts : la justification des postes de coût non pris en compte (dont en particulier celui des événements indésirables), l'estimation des volumes de ressources hospitalières et ambulatoires, les méthodes de valorisation et de calcul des coûts.

Un point particulièrement sensible est le recours fréquent à des avis d'experts pour quantifier les volumes et les fréquences de consommation de soins. Dans la quasi-totalité des cas, le rapport technique de l'évaluation ne détaille pas l'identité des experts, la méthode de recueil des avis et le traitement des réponses.

Comme pour les scores d'utilité, le niveau de la réserve finalement émise est très dépendant de l'impact attendu et des analyses mises en œuvre pour explorer cet impact. Des réserves importantes ont été émises lorsque la mésestimation des coûts était substantielle et favorable au produit (6 réserves importantes) ou lorsque la fiabilité des sources retenues ou des hypothèses pouvait être remise en cause (4 réserves importantes).

Pistes d'amélioration

Les pistes d'amélioration sont les mêmes que pour les scores d'utilité.

4.5. Résultats et analyses de sensibilité

► Le résultat de l'évaluation

Identification

Le résultat initial de l'évaluation de l'industriel, issu des choix méthodologiques dans l'analyse de référence, a été remis en cause lors de l'échange technique dans plus de la moitié des dossiers (18/30). Cela s'est traduit par une demande de modification de l'analyse de référence et de production d'un nouveau résultat. Une correction de la structure du modèle, qui peut être lourde, n'a été demandée que pour deux dossiers, toujours en précisant qu'une argumentation contraire de l'industriel était possible. Les modifications les plus fréquemment demandées concernent l'intégration de nouveaux comparateurs (6/18), le raccourcissement de la durée de simulation (6/18), et la redéfinition de la population d'analyse (4/18). Viennent ensuite : la modification d'une hypothèse (3/18), la mise à jour des prix (3/18), la modification de la population simulée (2/18), des coûts (2/18) ou des utilités (2/18).

Pour trois dossiers, il a été demandé lors de l'échange technique de refaire l'analyse selon une méthodologie adaptée à l'évaluation multi-options.

Tableau 5 : Détail des réserves majeures sur la présentation des résultats de l'évaluation économique

Résultats	Estimation d'un RDCR moyen sur la population d'analyse (Réserve majeure partielle) Méthode d'analyse non conforme à une évaluation multi-option Absence de frontière d'efficacité Interprétation fautive du résultat Résultats non transposables à une partie de la population d'analyse.
-----------	---

Explication

L'un des points sensibles de la présentation des résultats est la prise en compte de manière simultanée de tous les comparateurs retenus.

L'objectif principal de l'évaluation de l'efficacité est d'identifier les interventions situées sur la frontière d'efficacité. Ensuite seulement, le RDCR de chaque option peut être estimé, en rapportant le différentiel de coût au différentiel de résultat entre deux options successives sur la courbe d'efficacité. Lorsque l'évaluation introduit plusieurs comparateurs, la présentation des résultats deux-à-deux, sous la forme de RDCR estimés entre le produit et chacun de ses comparateurs n'est pas acceptable. L'absence d'une analyse adaptée à des comparateurs multiples a conduit la CEESP à émettre des réserves importantes ou majeures pour trois dossiers (2 réserves importantes et 1 réserve majeure).

Pistes d'amélioration

L'introduction de plusieurs comparateurs dans l'analyse se heurte aux limites de la métrique conventionnelle du RDCR, en particulier dans l'exploration de l'incertitude. La méthode du bénéfice net est privilégiée dans ces évaluations.

► Les analyses de sensibilité déterministes

Identification

L'échange technique est souvent l'occasion pour le SEESP de demander de nouvelles analyses de sensibilité déterministes (20/30), plus précisément sur : la durée de simulation (14/30), les méthodes d'extrapolation (13/30), les valeurs d'utilité (12/30), les données

cliniques introduites dans le modèle (12/30), les coûts (10/30), les taux d'actualisation (6/30), et les hypothèses de modélisation (6/30).

Des analyses de sensibilité sur les prix ont été demandées dans 9 dossiers, dont 5 portaient sur des variations du prix du produit étudié et 4 portaient sur des variations du prix d'un ou plusieurs comparateurs.

Suite à l'échange technique, les réserves importantes émises concernent principalement l'absence d'analyses sur des paramètres essentiels de la modélisation (12 réserves importantes) et l'absence d'analyses sur les prix (4 réserves importantes sur le produit évalué, 3 réserves importantes sur les comparateurs).

Explication

Les analyses de sensibilité déterministes sont essentielles pour identifier les paramètres dont la variation a le plus de poids dans l'évaluation de l'efficacité, et pour estimer l'impact attendu d'une éventuelle sous-estimation ou surestimation de ces paramètres. L'incertitude, ainsi analysée, participe à la qualification des réserves : une limite méthodologique ayant un impact fort sera davantage pondérée qu'une limite ayant un impact faible. À ce titre, le SEESP veille à ce que l'ensemble des paramètres incertains soient testés et est attentif au choix des bornes, des scénarios et des sources retenues pour construire ces analyses.

L'analyse de l'incertitude est souvent traitée de manière technique dans les dossiers, sans réelle interprétation des données quantitatives. Dans son analyse de l'incertitude des résultats présentés, le SEESP et la CEESP distinguent les différentes origines qui peuvent l'expliquer. L'analyse de l'incertitude associée aux choix structurants de l'évaluation soulève peu de réserves. Les analyses de sensibilité déterministes devraient explorer les deux sources d'incertitude qui ont le plus d'impact sur les résultats, à savoir l'incertitude inhérente aux choix de la modélisation (hypothèses de modélisation, méthodes d'extrapolation, sources de données) et l'incertitude sur la variabilité aléatoire des paramètres.

Pistes d'amélioration

Les limites associées aux analyses de sensibilité déterministes s'expliquent davantage par un manque de mise en perspective des analyses réalisées avec un objectif précis, que par des considérations techniques. S'interroger en amont sur le but de l'analyse de sensibilité devrait permettre de mieux cibler les paramètres à inclure dans l'analyse et d'interpréter les résultats en fonction de la nature de l'incertitude explorée, afin que les calculs présentés contribuent réellement à l'analyse de la robustesse de la modélisation.

► **Les analyses de sensibilité probabilistes**

Identification

Concernant l'analyse de sensibilité probabiliste (ASP), de nouvelles analyses sont plus rarement demandées que pour les analyses de sensibilité déterministes (7/30). Les questions visent à produire une analyse de sensibilité multi-option, à introduire une nouvelle variable ou à modifier un intervalle de confiance.

Les réserves émises sur les analyses de sensibilité probabilistes concernent l'absence de prise en compte de variables pertinentes (2 réserves importantes), la non pertinence ou la non justification des distributions (3 réserves importantes et 4 réserves mineures) et enfin le recours à une méthode inadaptée (4 réserves majeures).

Tableau 6 : Détail des réserves majeures sur les analyses de sensibilité de l'évaluation économique

Analyses de sensibilité	<p>Méthode d'analyse de sensibilité probabiliste non conforme à une évaluation multi-option</p> <p>Absence d'analyse de sensibilité probabiliste avec le bon comparateur</p> <p>Absence d'analyse de sensibilité probabiliste par sous-population</p> <p>Absence d'analyse de sensibilité probabiliste pour une sous-population (R^{++} partielle). Les RDCR des autres sous-populations sont recevables.</p> <p>Absence d'analyse de sensibilité probabiliste pour certaines sous-populations (R^{++} partielle). Les RDCR des autres sous-populations sont recevables.</p>
-------------------------	--

Explication

Les analyses de sensibilités probabilistes permettent d'explorer la stabilité des conclusions, considérant la variabilité aléatoire des estimations ponctuelles des paramètres du modèle. En faisant varier tous les paramètres en même temps, elle complète les analyses de sensibilité déterministes réalisées pour chaque paramètre pris individuellement.

L'information produite est fondamentale pour renseigner le CEPS sur le degré d'incertitude associé à l'évaluation. Les nuages de points renseignent sur la dispersion attendue des résultats et les courbes d'acceptabilité permettent d'appréhender le degré de confiance que l'on peut accorder à une conclusion d'efficacité, en fonction de la disposition à payer de la société. L'analyse de sensibilité probabiliste fondée sur la métrique du RDCR ne peut produire que des courbes d'acceptabilité comparant deux options. Lorsque plus de deux options sont comparées, seules les courbes d'acceptabilité multi-options fondées sur la métrique du bénéfice net peuvent documenter l'incertitude. Compte tenu de l'enjeu, les réserves formulées sur le choix inadapté de la méthode sont le plus souvent importantes ou majeures. Au vu des derniers dossiers soumis, il semble cependant que les courbes d'acceptabilité multi-options soient de plus en plus utilisées.

Pistes d'amélioration

La principale lacune de l'analyse probabiliste dans les dossiers soumis à la CEESP est l'absence d'interprétation des résultats en termes d'incertitude paramétrique. Trop souvent, l'industriel se contente de rappeler les données quantitatives lisibles sur les courbes d'acceptabilité. Une façon d'améliorer l'information transmise serait de prendre réellement position sur l'ampleur de l'incertitude.

5. DISCUSSION

L'analyse chronologique des 30 premiers avis émis par la CEESP permet d'observer une amélioration de la qualité de l'information transmise au CEPS, qui s'explique par une meilleure qualité technique des évaluations réalisées et par une évolution de la rédaction de la conclusion de la commission, davantage centrée sur l'interprétation des résultats de l'évaluation.

Participant à ce constat, l'échange technique est devenu une étape clé dans le processus d'expertise des évaluations transmises à la CEESP. Sur la période, le nombre de questions techniques a beaucoup augmenté, avec une plus grande précision des questions posées, afin d'aider l'industriel à corriger les limites constatées en première analyse. L'objectif de cet échange est de réduire l'incertitude. L'analyse conjointe des questions posées et des réserves finalement émises montre que les réponses de l'industriel permettent de clarifier de manière substantielle certaines dimensions, telles que la gestion de la dimension temporelle.

Afin de poursuivre dans cette direction, la HAS a identifié trois axes visant à améliorer la qualité de l'information disponible sur l'efficacité des produits qu'elle est chargée d'évaluer : clarifier ses attentes auprès des industriels, actualiser le guide méthodologique pour l'évaluation économique [4] et cibler des problématiques de recherche méthodologique.

Des attentes à clarifier

Un grand nombre de questions techniques ont pour objet la justification de choix méthodologiques ou la clarification de la méthode appliquée. Ces questions traduisent davantage des insuffisances de présentation que de réelles limites méthodologiques.

A partir de ce constat, le SEESP a réalisé un format de rédaction des rapports techniques destiné aux industriels. Ce format explicite de manière précise les éléments de dossiers attendus par la HAS, ce qui devrait favoriser l'obtention des informations nécessaires à l'expertise, dès la soumission de l'évaluation. Ce format a été soumis à une consultation publique au cours de l'été 2015, et sa publication est prévue au premier trimestre 2016.

En particulier, le format rappelle la nécessité d'une interprétation des résultats et de l'incertitude. En effet, au travers des 30 dossiers analysés, la HAS a pu constater que l'interprétation des résultats quantitatifs de la modélisation est souvent trop peu développée dans les dossiers soumis, au profit d'une technicisation des modèles et d'une multiplication des résultats produits.

La production d'une discussion apportant une interprétation utile à la décision est une attente forte. Cela implique une prise de position claire, de la part de l'industriel et de la HAS, qui ne peut se limiter à un simple rappel des résultats quantitatifs. Interpréter l'évaluation implique d'être capable d'estimer le degré de confiance qu'il est possible d'accorder au résultat et à sa transposition en pratique courante.

Dépasser la dimension technique, pour donner du sens à l'évaluation réalisée, constitue probablement le plus grand défi pour tous les acteurs concernés.

Des principes méthodologiques à actualiser

L'analyse des questions techniques et des réserves permet d'ores et déjà d'identifier certaines dimensions de l'évaluation économique qui devront être introduites ou réexaminées dans le cadre de l'actualisation du guide méthodologique de la HAS [4]. Une nouvelle version du guide HAS devrait être soumise à la consultation publique fin 2016.

Le retour d'expérience a permis d'identifier deux problématiques nouvelles par rapport aux dimensions identifiées en 2011.

- Premièrement, le principe de la cohérence entre la revendication, l'objectif de l'évaluation déposée et les choix méthodologiques devrait être affirmé. La production d'un RDCR n'est pas un objectif en soi. Un des enjeux pour les industriels est, d'une part, de formuler un objectif qui a du sens au regard de leur revendication en termes d'inscription (primo-inscription, réévaluation ou réinscription, extension) ; d'autre part, de définir les choix structurants de l'évaluation en fonction de cet objectif (notamment le choix des comparateurs et la définition de la population d'analyse).
- Deuxièmement, l'essor d'évaluations économiques introduisant de multiples comparateurs devrait conduire à développer la métrique du bénéfice net, plus adaptée à ces évaluations que la métrique du RDCR.

D'autres dimensions devraient faire l'objet d'une reformulation. Par exemple, les définitions des notions de perspective, de population d'analyse ou de population simulée, devraient être clarifiées.

Concernant les autres dimensions de l'évaluation, par exemple la modélisation ou les méthodes de méta-analyse, il s'agit davantage de tenir compte de l'évolution des connaissances scientifiques, avec une évolution souhaitée par les industriels vers un guide

plus opérationnel sur le plan technique. En particulier, ce premier bilan permet de dresser le constat d'une incertitude encore forte sur les résultats des modélisations. La caractérisation de cette incertitude est un enjeu majeur de l'évaluation économique et de sa capacité à éclairer la décision en matière de fixation des prix. Cela implique une amélioration de la qualité des méthodes de validation et d'estimation de la sensibilité des résultats aux choix méthodologiques et à la variabilité des paramètres.

Enfin, parmi les dossiers déposés par les industriels, une minorité comportait une analyse d'impact budgétaire. La CEESP a constaté à plusieurs reprises qu'une telle analyse permettrait de mieux documenter le critère économique. Un travail méthodologique sur cette méthode d'évaluation est en cours à la HAS.

Des défis méthodologiques à relever

Au-delà de ce travail d'actualisation des principes méthodologiques de la HAS, certaines difficultés constatées soulèvent des enjeux méthodologiques qui appellent des propositions nouvelles. Trois axes ont été identifiés, sans avoir l'ambition d'épuiser l'ensemble des possibilités.

Premièrement, les limites concernant les données d'efficacité renvoient aux choix méthodologiques établis lors de la réalisation des essais cliniques. La qualité de la preuve sur l'apport des produits est une limite régulièrement soulevée par les commissions médico-techniques de la HAS (CT, CNEDIMST). Dans un contexte où les demandes d'expertise précoces tendent à se multiplier sur la base d'essais de phase II, l'inscription initiale s'apparente à un premier jalon dans l'évaluation du produit, dont le moment fort pourrait devenir le renouvellement d'inscription. Cela implique que des études post-inscription de qualité soient mises en place par les industriels, pour documenter les paramètres les plus incertains de l'évaluation médico-économique déposée au moment de la primo-inscription.

Deuxièmement, l'identification de tous les comparateurs pertinents reste une exigence pour interpréter l'évaluation présentée. Les réserves émises montrent que ce principe de transparence n'est pas encore suffisamment respecté. En revanche, leur introduction dans l'évaluation peut poser de réelles difficultés techniques, en particulier en l'absence de données cliniques ou de prix. Une réflexion et des développements méthodologiques sont donc nécessaires sur la meilleure façon de prendre en compte ces comparateurs.

Troisièmement, les limites récurrentes identifiées concernant la qualité des scores d'utilité ou des coûts introduits dans l'évaluation inciteraient à développer des études afin de produire des données françaises standardisées. La disponibilité de telles données auraient l'avantage de réduire l'hétérogénéité de l'incertitude dans les évaluations.

6. Conclusion

Au 31 juillet 2015, la CEESP a rendu 30 avis d'efficience. L'avis de la CEESP offre un outil complémentaire au CEPS pour négocier les prix des produits de santé [8]. Dans chaque situation, la CEESP tente d'apporter un éclairage sur les éléments dont elle dispose pour caractériser l'efficience, bien qu'il n'existe pas de valeur de référence [9].

A ce jour, sur les 30 avis rendus, 13 produits ont vu leur prix publié au Journal Officiel, ce qui ne permet pas d'estimer un impact quantitatif des avis d'efficience sur les niveaux de prix obtenus.

- L'avis permet d'indiquer au CEPS les produits pour lesquels l'efficience n'est pas démontrée, soit parce qu'un choix méthodologique n'est pas acceptable, soit parce que l'accumulation de l'incertitude rend le RDCR estimé par l'industriel trop fragile pour être interprétable.

- Dans les 20 dossiers pour lesquels une évaluation du RDCR est disponible, l'avis de la CEESP rend explicite le surcoût d'un produit de santé et les résultats de santé supplémentaires qu'il génère.
- A partir des données produites par l'industriel, l'avis fournit au CEPS une interprétation des résultats et de l'incertitude qui l'entoure, de manière à estimer la fiabilité d'une transposition des résultats en pratique courante. Cette évaluation initiale permet au CEPS d'identifier les produits pour lesquels l'évaluation initiale est susceptible d'être fortement révisée. Dans un contexte où les demandes d'expertise précoces tendent à se multiplier sur la base d'essais cliniques de phase II, cette évaluation renforce l'assise d'accords de prix visant à limiter les effets d'une inscription précoce fondée sur des bases fragiles.

Enfin, l'identification des limites méthodologiques des évaluations soumises permet d'orienter le décideur public quant aux demandes complémentaires à solliciter lors du renouvellement d'inscription. En décrivant et en qualifiant l'incertitude sur le résultat de l'évaluation, l'avis de la CEESP permet d'avoir une idée de l'amplitude de la différence entre l'efficacité théorique estimée et l'efficacité anticipée en pratique, ce qui apporte des arguments pour fonder les accords de prix, ou d'autres types de clauses, type accords de performance.

Références

- [1] République française. Décret n° 2012-1116 du 2 octobre 2012 relatif aux missions médico-économiques de la Haute Autorité de santé. JORF n°0231 du 4 octobre 2012. p 15522.
- [2] République française. Loi n°2014-1554 du 22 décembre 2014 de financement de la sécurité sociale pour 2015. JORF n°0297 du 24 décembre 2014. p 21748.
- [3] Décret n° 2012-1116 du 2 octobre 2012 relatif aux missions médico-économiques de la Haute Autorité de santé JORF n°0231 du 4 octobre 2012 page 15522 texte n° 8.
- [4] HAS. Choix méthodologiques pour l'évaluation économique à la HAS. Octobre 2011, 82 pages. www.has-sante.fr
- [5] HAS. Rapport d'activité pour 2014. Juin 2015, 144 pages. www.has-sante.fr
- [6] Jeantet M., Lopez A., Destais N. Evaluation médico-économique en santé. Rapport 2014-066R, décembre 2014, 203 pages.
- [7] Haute Autorité de santé. Format de l'avis d'efficacité. Paris : Haute Autorité de Santé ; Juillet 2013. 24 p. Disponible sur www.has-sante.fr
- [8] Comité économique des produits de santé. Rapport d'activité 2014/2015. Septembre 2015. 195 p. Disponible sur <http://www.sante.gouv.fr/les-activites-du-ceps.html>
- [9] HAS. Valeurs de référence pour l'évaluation économique – Revue de littérature. Décembre 2014, 76 pages. www.has-sante.fr