



Synthèse des données probantes et évaluation médico-économique

Les deux doigts d'une même main

Quantitative systematic review and health technology assessment

Two peas in a pod

R. Launois

Directeur Réseau d'Évaluation en Économie de la Santé (REES France),
28, rue d'Assas - 75006 PARIS, launois.reesfrance@wanadoo.fr, www.rees-france.com

Mots clés : synthèse quantitative, mélange de comparaison de traitement.

Keywords: quantitative synthesis, mix treatment comparisons.

Résumé

Si les essais cliniques randomisés permettent d'estimer la quantité d'effet produite par un traitement au regard d'un autre, il est fondamental de replacer les résultats obtenus par un traitement innovant par rapport à l'ensemble des traitements existants dans l'indication. C'est de cette confrontation élargie qu'il sera possible de conclure à l'existence ou non d'un véritable progrès thérapeutique. Instruments incontournables de la connaissance scientifique, les essais randomisés ne peuvent à eux seuls éclairer la décision politique et ne sauraient être le vecteur de l'évaluation des technologies de santé. Deux types d'outils semblent aujourd'hui à même de pallier les éventuelles insuffisances des schémas expérimentaux : les méta-analyses bayésiennes et les analyses décisionnelles stochastiques. Cet article se propose d'étudier ces outils de deuxième génération et de les codifier pour faciliter leur mise en œuvre dans des domaines où il existe de véritables enjeux de santé publique : prévention et dépistage du cancer, traitements de première et de deuxième ligne des cancers métastatiques ou non du sein, du poumon, du colon etc.

Summary

A common feature of available evidence is the absence of head-to-head trials comparing all relevant options being considered. Instead, there is a series of trials making different pair-wise comparisons among the interventions of interest. Health authorities are interested in comparing a new treatment to several alternatives. The "historical approach", extract arms from different clinical trials and plug them into a Markov model taking for granted that the populations in the trials are comparable. This approach, clearly "breaking" the randomisation, leads to bias and is more and more considered as inappropriate. Mixed treatment comparisons model allows dealing with that kind of evidence structure, where both direct and indirect evidence may be used to estimate the parameters.

Lorsqu'un produit reçoit son autorisation de mise sur le marché, il n'a pas dans tous les cas fait l'objet d'une comparaison avec un autre traitement actif dans le cadre d'un essai clinique randomisé en double aveugle. Or c'est précisément à ce moment que le décideur doit en fixer le prix et préciser les modalités de sa prise en charge. L'enjeu essentiel pour le décideur est d'estimer la plus value thérapeutique de l'innovation en comparant ses performances avec celles qui sont obtenues par chacun des traitements couramment prescrits dans la même indication. Une seule confrontation de deux molécules en tête à tête ne

permet pas de se prononcer sur le bénéfice attendu par rapport à la palette des traitements utilisés. Il convient de juger l'impact thérapeutique différentiel d'un nouveau traitement à l'aune de celui qui peut être obtenu par des comparateurs déjà présents sur le marché.

Pour estimer l'amélioration du service médical rendu l'analyste est amené à confronter les résultats d'essais multiples dont les designs, les populations, les critères de jugement et les comparateurs sont bien souvent différents. Jusqu'à présent la synthèse des informations pertinentes relevait d'une « analyse statistique raisonnée » dont

la démarche n'était pas sans rappeler le rôle et l'importance du jugement clinique dans la décision médicale. L'analyste cherchait à réaliser les combinaisons optimales à partir des données médicales disponibles pour prévoir les trajectoires de vie concevables et chiffrer leurs conséquences humaines, sociales et financières. Une revue non exhaustive de la littérature permettait d'alimenter le modèle. Les options thérapeutiques testées dans des essais différents, réalisés sur des populations comparables mais non strictement superposables, étaient rapprochées après validation par un comité d'experts. Le choix des essais demeurait entaché de subjectivité et de toute façon les ran-

domisations étaient systématiquement brisées. Aujourd'hui on ne peut plus construire les modèles sur de telles bases.

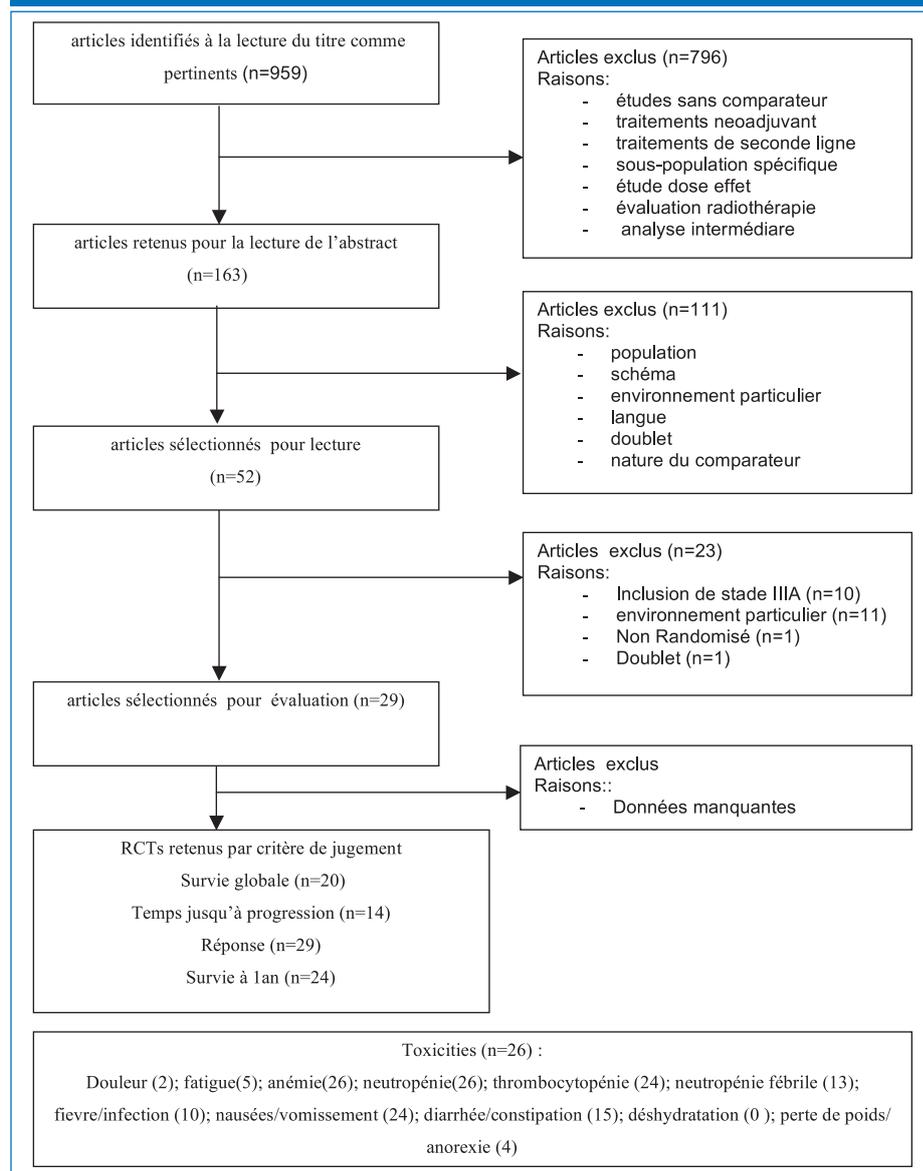
Tout modèle exige désormais une synthèse robuste des données probantes très en amont de sa construction. L'ère des panoramas brillants mais narratifs s'achève. Celle des revues systématiques de la littérature et des synthèses quantitatives des données probantes commence.

Bien qu'il règne une certaine confusion au sujet de la signification de ces deux termes, il est utile d'en distinguer très clairement le contenu. La première méthode permet de recueillir de manière systématique et exhaustive l'ensemble des informations disponibles sur les différents traitements étudiés sans combiner quantitativement le fruit de ses recherches. La démarche est méthodique et systématique sans qu'un effet commun à l'ensemble du faisceau de preuves réunies ne soit estimé. On parle dans ce cas de synthèse qualitative des données probantes (1, 2, 3, 4, 5, 6). La seconde permet d'estimer de manière quantitative la taille des effets des traitements ou d'un mode d'organisation sur l'efficacité des prises en charge observée dans les différentes études publiées. Cette approche se réfère explicitement aux méta-analyses (7, 8) et aux méthodes statistiques de mélange de comparaisons de traitements (9, 10).

Synthèse qualitative des données probantes

Les objectifs explicites de l'étude doivent être clairement établis. De façon plus précise la question doit être décomposée en fonction des critères PICOT. Un acronyme dont les composantes désignent: les caractéristiques des patients et des professionnels (P), auxquelles on s'intéresse. Pour les premiers il s'agira de spécifier le stade de la maladie, la ligne de traitement, l'état général. Pour les seconds de préciser dans quel(s) cadre(s) ils exercent leurs activités ; la nature de l'intervention « I » ; le comparateur choisi « C » ; les « outcomes » (O) c.à.d. les résultats clés retenus comme critères de jugement autrement dit les mesures quantitatives mises en oeuvre pour estimer l'efficacité et la sécurité, l'horizon temporel (T) retenu pour évaluer les résultats.

Figure 1 : Diagramme de sélection des traitements de première ligne dans le cancer bronchique non à petites cellules

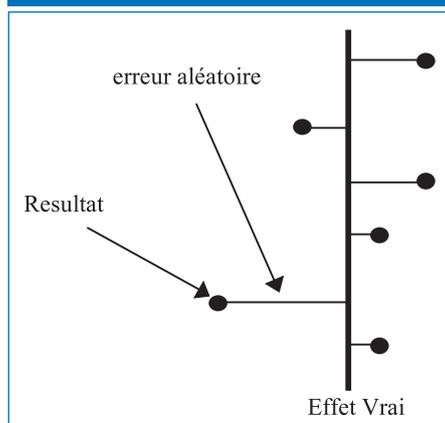


tats. Cette quantité d'effet est appréhendée à travers un nombre limité d'indices thérapeutiques, indices parce qu'il s'agit toujours d'une transformation des résultats bruts, destinés à faire apparaître plus clairement les avantages que présente un traitement par rapport à un autre (11, 12, 13, 14).

C'est aux auteurs de la synthèse qualitative de la littérature de décider quel sera le niveau de preuve sur lequel ils s'appuieront pour tirer leur conclusions au regard des indices précédemment décrits. Les exigences en ce domaine sont plus ou

moins restrictives. La plupart des revues systématiques du groupe Cochrane reposent exclusivement sur des essais randomisés ; mais ils leur arrivent dans certains cas de s'appuyer sur d'autres schémas d'études. Par exemple, un grand nombre des revues qui ont été faites par le Cochrane effective practice and organisation of care (EPOC) intègre les résultats issus d'études quasi-expérimentales avec comparateur. Dans tous les cas il faut vérifier quel type d'études aura le niveau de preuve suffisant et le plus adapté à la question posée. La hiérarchie des preuves adoptée doit être précisée.

Figure II : modèle à effet fixe



La sélection des articles susceptibles de faire l'objet d'une étude approfondie se fait à travers plusieurs filtres et le processus de sélection doit être décrit de façon précise. La simple lecture du titre ou de l'abstract des articles recensés à l'issue de la recherche électronique permet de découvrir si chacun d'entre eux répond ou non aux critères d'éligibilité retenus. Sur la base des informations disponibles, il peut apparaître clairement que l'article en question ne permettra pas de répondre à la question posée ; Auquel cas il doit être rejeté. Lorsqu'un doute subsiste l'article devra être commandé et lu dans son intégralité. Après lecture, il peut encore être écarté s'il ne remplit pas les conditions d'éligibilité retenues. Dans le cas contraire, il fera l'objet d'une analyse structurée conforme à la grille de lecture qui aura été définie ex ante : objectif de l'étude, type de protocole, intervention étudiée, critères de jugement, population source, facteurs de confusion et biais, analyses statistiques, conclusions des auteurs.

Un exemple du processus de sélection des études dans une indication sur laquelle nous travaillons actuellement est présenté dans la figure I.

Meta analyse et synthèse quantitative

L'efficacité des traitements médicaux est généralement évaluée par comparaison directe avec un ou plusieurs comparateurs à travers des essais randomisés. Si l'essai randomisé est l'instrument incontournable de l'évaluation de l'efficacité relative des pro-

duits, il n'est pas le bon vecteur pour procéder à l'évaluation des technologies médicales :

- Les essais randomisés se réduisent souvent à une confrontation de 2 traitements en tête-à-tête. Ils ne permettent pas d'identifier les avantages et les inconvénients d'un nouveau traitement par rapport à l'ensemble de l'arsenal thérapeutique existant.
- Les essais randomisés négligent les répercussions financières et psychologiques des traitements mis en œuvre. Ils reposent sur une approche partielle du problème à traiter.

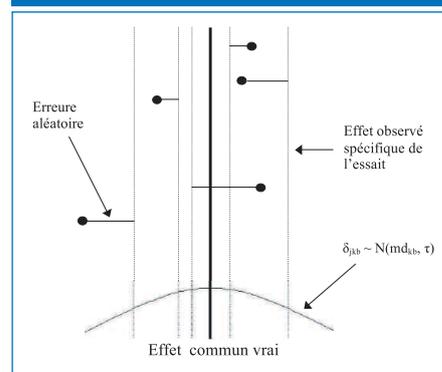
Les méthodes statistiques quantitatives sont alors d'un précieux secours pour déterminer l'ampleur moyenne de la quantité d'effet relevée dans des études de même nature qui organisent la confrontation des mêmes options thérapeutiques en tête-à-tête ou pour situer la place d'un traitement dans l'ensemble de l'arsenal thérapeutique existant pour une indication déterminée. Dans le premier cas ce seront les techniques de la méta-analyse qui seront mises en œuvre. Dans le second on aura recours aux modèles statistiques de mélange de comparaisons de traitements.

Les méta-analyses permettent d'augmenter la puissance des essais cliniques et de synthétiser leurs résultats de façon quantitative exclusivement lorsqu'il s'agit des mêmes comparaisons par paires.

Dans un modèle à effets fixes (15), on suppose que chaque étude estime le même paramètre sous-jacent. Autrement dit, la variabilité observée entre les effets estimés dans chaque étude ne provient que de la précision avec laquelle les effets sont mesurés. Dans ce cas, l'effet global du traitement correspond à la moyenne pondérée des effets estimés dans chaque étude, les poids associés étant égaux à la précision avec laquelle les effets sont mesurés.

De nombreux auteurs ont montré que l'hypothèse soutenant les modèles à effets fixes n'est pas réaliste. En effet, les études portant sur le même sujet ne sont pas des répliques les unes des autres, elles ne suivent pas forcément le même schéma d'étude

Figure III : Modèle à effet aléatoire

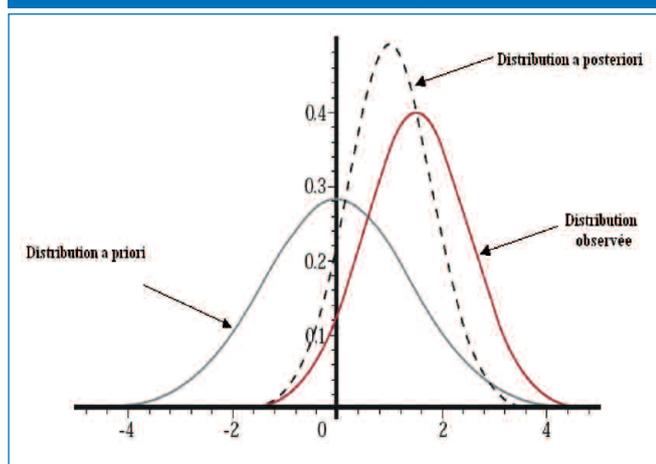


et sont conduites différemment. Cela a donc des conséquences sur la mesure du critère d'efficacité du traitement.

Les modèles à effets aléatoires (16) permettent de tenir compte de cette variabilité inter-études. Dans ces modèles, on abandonne l'hypothèse selon laquelle chaque étude estime exactement le même effet. En revanche, on suppose que chaque étude estime un effet différent, mais que tous ces effets sont issus d'une même distribution d'effets. Il est commode de représenter les incertitudes portant sur la nature incertaine des effets par une distribution de probabilité ; il suffit alors d'estimer les valeurs d'un petit nombre de paramètres pour définir toute la distribution de probabilité. Bien que les méta analyses utilisent les résultats voire les données individuelles des essais cliniques, leurs conclusions ne découlent pas directement d'un plan d'expérience, la présence de biais est donc toujours possible. Mais surtout elles souffrent des mêmes limitations que les essais dont elles sont issues : comparaison impossible de toutes les options thérapeutiques ; vision tronquée du génie évolutif de la maladie, négation des réalités épidémiologiques et institutionnelles locales, scotomisation d'éléments décisifs pour la prise de décision politique : qualité de vie, trajectoires et contacts, diversité des modes de prises en charge.

La description ci-dessus des modèles à effets fixes et à effets aléatoires se place dans un contexte classique, fréquentiste, c'est-à-dire que les paramètres sont supposés fixes et inconnus (θ et μ par exemple). Dans un paradigme bayésien, au contraire, les para-

Figure IV : Modèle Bayésien



mètres sont supposés aléatoires, ce qui permet d'associer une distribution de probabilité sur chacun d'entre eux. Alors que l'école classique fonde le jugement sur une succession d'échantillon virtuels dont la loi statistique est conditionnelle à des paramètres fixes mais inconnus, l'école bayésienne considère que les paramètres sont des variables aléatoires. Il y a donc un sens à considérer leur distribution probabiliste et à les revoir à la lumière des nouvelles informations. Cette propriété ouvre des possibilités supplémentaires en terme de modélisation. Il devient possible de spécifier des distributions a priori pour chaque paramètre, fondées sur des connaissances antérieures. En effet, avant de conduire une étude, un chercheur peut avoir formé une intuition a priori concernant l'option thérapeutique étudiée. Cette intuition peut être fondée sur sa propre expérience ou sur des études antérieures menées sur le même thème.

Les modèles bayésiens permettent de combiner ces intuitions ou connaissances a priori avec les données issues de l'étude menée (grâce au théorème de Bayes). Les résultats reflètent alors l'actualisation de l'état des connaissances sur l'effet de la stratégie considérée.

Mélange de comparaisons de traitements

L'appréciation de l'amélioration du service médical attendu doit se faire par rapport à l'efficacité et à la sécurité des traitements existants. Or les essais présentés en commission de la transparence n'ont

pas toujours le bon comparateur. S'ils ont été mis en place en Amérique du Nord, il peut s'agir d'un simple placebo ou des traitements usuels de référence sans que les résultats spécifiques de chacun d'entre eux soient isolés. Lorsque le ou les comparateurs sont des traitements actifs, leur nombre dépasse rarement deux ou trois. Le plus souvent il s'agit de comparaisons en tête en tête. Situer l'innovation par rapport à l'arsenal thérapeutique existant n'est donc possible qu'en ayant recours à des comparaisons indirectes.

Trois méthodes permettent d'y procéder : les « raboussages naïfs », les comparaisons indirectes ajustées sur un dénominateur commun, les mélanges de comparaisons directes et indirectes de traitements (MTC).

Dans le premier cas les résultats obtenus sous différents traitements dans divers essais sont intellectuellement contrastés les uns par rapport aux autres ou explicitement et quantitativement intégrés dans un arbre de décision en prenant les valeurs absolues de survenue des événements comme probabilités. Une telle synthèse des données probantes est hautement contestable. Le rapprochement des résultats absolus observés sur différents bras de traitement empruntés à des études distinctes est fatalement biaisé puisque les observations peuvent avoir été obtenues sur des populations très différentes. Ces comparaisons intuitives et non scientifiques étaient il y a dix ans les seules possibles ; elles prétendaient mêmes être « evidence based » puisqu'elles reposaient sur des données issues d'essais randomisés. Ce raboussage artificiel a très certainement été à l'origine de la méfiance qu'ont inspirée les modélisations. Le problème est aujourd'hui résolu grâce aux techniques statistiques bayésiennes.

En 1997 Butcher et Guayatt (17) jettent les bases d'une solution en proposant un modèle statistique de comparaisons indirectes ajustées sur un dénominateur commun. La méthode s'appuie sur le principe

de transitivité pour rapprocher quantitativement les résultats des traitements pour lesquels il n'existe pas de données directement comparatives dans les essais cliniques. C'est l'effet différentiel des traitements qui est pris en compte. On l'estime par rapport à un traitement de référence commun aux différents essais sélectionnés. Les effets absolus des différents traitements dont l'ampleur peut être liée à la spécificité des populations recrutées ne sont pas pris en considération. Le raisonnement en quantité d'effet différentielle assure une relative comparabilité des groupes et réduit d'autant le risque de biais. Dans les modèles de mélange de comparaisons de traitements, on travaille toujours sur les différences d'effet par rapport à un comparateur commun mais cette fois le comparateur commun est choisi par l'analyste lui-même puis utilisé comme numéraire dans des rapports de cotes dont le numérateur et le dénominateur sont définis au regard du même traitement de référence. Il suffit alors de prendre le logarithme népérien du rapport de cote pour transformer le quotient en simple différence de logarithmes de cotes, ce qui permet de faire jouer les propriétés de transitivité. L'ensemble des essais peut être mobilisé, en comparant les traitements soit directement (en face à face) ou soit indirectement (par transitivité) sans « briser la randomisation » effectuée dans les essais de départ. L'outil statistique mis en œuvre est bayésien. Alors que l'école classique fonde le jugement sur une succession d'échantillon virtuels dont la loi statistique est conditionnelle à des paramètres fixes mais inconnus, l'école bayésienne considère que les paramètres sont des variables aléatoires. Il y a donc un sens à considérer leur distribution probabiliste à priori et à les revoir à la lumière des nouvelles informations; Plusieurs types de lois a priori peuvent être utilisées. Lorsque aucune intuition à priori sur les paramètres n'est possible, des lois a priori dites non informatives sont employées.

L'avantage de la méthode est double. L'inférence concernant l'efficacité relative de deux traitements s'en trouve renforcée puisque les données probantes incluent à fois des preuves directes et indirectes. L'inférence simultanée sur l'efficacité relative de l'en-



semble des traitements devient possible, et la place d'un nouveau traitement dans l'ensemble des stratégies thérapeutiques existantes peut être déterminée de façon rigoureuse. Les mélanges de comparaisons de traitement offrent la solution qui permet d'associer inférence statistique, synthèse des données probantes et prise de décisions.

Illustration

Les 4 stratégies utilisées en 2^{ème} ligne de chimiothérapie dans le cancer bronchique non à petites cellules (CBNPC) : Docetaxel toute les trois semaines, Pemetrexed toute les trois semaines, Docetaxel hebdomadaire 35 mg/m². et Docetaxel hebdomadaire 36 mg/m² n'ont jamais été comparées en tête à tête au sein d'un seul et même essai. Les données d'efficacité (durée médiane de survie et temps médian jusqu'à progression) et de tolérance sont partagées en 3 essais (18, 19, 20, 21), comparant les stratégies 2 à 2.

Il faut donc synthétiser de manière appropriée l'information issue de ces 3 essais pour obtenir des données cohérentes. Les modèles de mélange de comparaisons de traitements (Mixed Treatment Comparisons, MTC) permettent de réaliser des comparaisons indirectes non biaisées.

Ils ont donc été employés pour synthétiser l'information sur 3 types de critères :

- La durée médiane de survie.
- Le temps médian jusqu'à progression.
- Les toxicités.

La durée médiane de survie et le temps médian jusqu'à progression ont été modélisés de la même manière. Les hypothèses sous-jacen-

Tableau 1 : Pourcentage de patients souffrant de neutropénie fébrile selon l'essai

	Pemetrexed tout les 21 j	Docetaxel tout les 21 j	Docetaxel hebdo 35 mg/m ² .	Docetaxel hebdo 36 mg/m ² .
Hannah	1,9 %	12,7 %		
Schuette		2,0 %	1,0 %	
Camps		7,8 %		0,8 %

Tableau 2 : Résultats du modèle MTC en termes de pourcentage de patients souffrant de neutropénie fébrile

Traitement	Moyenne	Percentile 2,5 %	Percentile 97,5 %
Docetaxel	0,055	0,028	0,086
Pemetrexed	0,007	0,000	0,087
Docetaxel W35	0,021	0,000	0,352
Docetaxel W36	0,004	0,000	0,077

tes permettant les calculs sont complexes. Pour simplifier, on a supposé que la survie suivait une loi exponentielle, tandis que les valeurs médianes suivaient une loi lognormale.

Il est beaucoup plus simple d'illustrer le principe des MTC en utilisant les toxicités. L'exemple qui suit concerne donc les neutropénies fébriles. Initialement, les 3 essais donnent les pourcentages de patients souffrant de neutropénie fébrile suivants :

Les comparaisons indirectes s'avèrent utiles lorsque les traitements d'intérêt n'ont jamais été comparés de manière directe, mais qu'il existe d'autres traitements auxquels ils ont été comparés directement. Dans le cas du CBNPC par exemple, l'essai de Hannah compare les traitements Pemetrexed et Docetaxel 3 semaines et l'essai de Schuette compare Docetaxel 3 semaines et Docetaxel hebdomadaire 35 mg/m². Alors, par transitivité, on peut obtenir une comparaison indirecte du Pemetrexed et du Docetaxel hebdomadaire 35 mg/m², comme le montre le graphique suivant.

Le modèle de mélange de comparaisons de traitements permet ainsi de mobiliser l'ensemble des essais disponibles, qu'ils permettent de comparer les traitements directement (en face à face) ou indirectement (par transitivité).

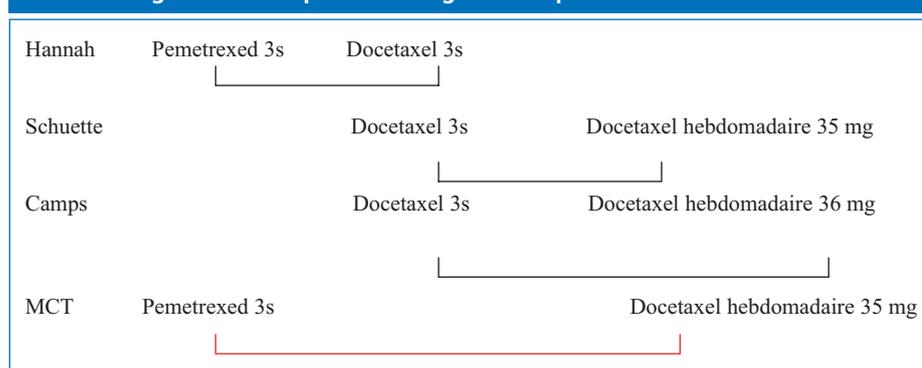
Les MCT peuvent être modélisées par des modèles hiérarchiques bayésiens. Les modèles bayésiens permettent de combiner une connaissance a priori sur les paramètres aux données disponibles, représentant l'acquis de connaissances nouvelles sur ces paramètres. Le logiciel WinBUGS permet de construire ce type de modèle et estime les paramètres en construisant des chaînes de Markov par la méthode de Monte Carlo.

En utilisant les principes de calculs décrits précédemment, le modèle aboutit aux résultats suivants :

La stratégie la plus toxique en terme de neutropénie fébrile est donc le Docetaxel. Le Pemetrexed n'apparaît qu'en 3^{ème} position.

Un des inconvénients des comparaisons indirectes est qu'elles sont plus susceptibles d'être biaisées que les comparaisons directes. En fait, cela dépend en grande partie de la façon dont les études sont mises en commun. La mise bout à bout des résultats obtenus sur les différents bras dans des études différentes conduit inmanquablement à des résultats biaisés. En modélisation MCT, le paramètre d'intérêt reflète la différence d'effet entre traitements plutôt que l'effet des traitements eux-mêmes, ce qui permet d'échapper à cet écueil. Une technique qui, par ailleurs, ne brise pas la

Figure V : Principe du mélange de comparaison de traitement



randomisation effectuée dans les essais de départ.

Une analyse décisionnelle renouvelée

Sachant que dans le domaine de la santé les décisions réversibles sont plutôt rares, la seule vraie question est de savoir comment reconnaître une bonne décision d'une mauvaise. Toute argumentation en faveur de telle ou telle option ne peut faire l'impasse sur le risque global qui lui est associé. Toute décision médicale ou tout choix politique doit donc être fondé sur une estimation quantitative des conséquences dommageables ou bénéfiques entraînées par chacune des options thérapeutiques disponibles. L'analyse décisionnelle permet de prendre de façon systématique et simultanée l'ensemble des enjeux médicaux et socio-économiques sous-jacents à la prise de décision.

Trois éléments caractérisent un choix thérapeutique : les traitements disponibles ; les probabilités de survenue des événements conditionnellement au traitement choisi et leurs lois de distributions ; les conséquences dommageables ou bénéfiques qui en résultent et leurs contre-valeurs monétarisées.

A chaque paramètre clinique utilisé dans l'arborescence on associe une valeur. Dans le cadre d'une analyse déterministe c'est le risque absolu de survenue de l'événement conditionnellement au traitement choisi qui est utilisé en se fondant sur le résultat des essais randomisés publiés. Pour chaque modalité de traitement envisagé, il est calculé une espérance globale d'efficacité et une espérance globale de coût du traitement. Cette estimation ponctuelle du risque réduit l'information en masquant la variabilité des résultats attendus par le biais d'une moyenne et dissimule les incertitudes liées aux inévitables fluctuations d'échantillonnage.

Ces deux notions de variabilité et d'incertitude sont essentielles dans l'évaluation quantitative du risque. La variabilité se réfère à une caractéristique intrinsèque du phénomène étudié à laquelle il est impossible d'échapper. Cette variabilité par essence

est une donnée naturelle irréductible. On observera toujours des différences entre les groupes ou à l'intérieur d'un groupe. L'augmentation du nombre de données ne remettra jamais d'échapper à la réalité de ces phénomènes aléatoires. L'incertitude due à l'ignorance est au contraire réductible par apport de données nouvelles ce qui permet d'orienter les activités de recherches ultérieures et ouvre la voie aux applications de la théorie de la valeur escomptée de l'information. La nature de l'incertitude peut être de diverses origines : choix du modèle, représentativité des données empiriques disponibles, taille finie de l'échantillon à partir duquel ont été estimés les paramètres de la population.

L'analyse probabiliste du risque encore appelée analyse stochastique permet de prendre en compte simultanément la variabilité du risque et l'incertitude liée à son estimation. La démarche est simple dans son principe. A chaque variable du modèle est associée non plus une valeur moyenne ponctuelle mais une distribution de valeurs possibles associées à cette valeur moyenne ; Autrement dit, à chaque probabilité est associée une distribution de probabilité. La principale difficulté provient du choix de ces lois de densité pour chacune des variables identifiées. Toute loi de densité est caractérisée par sa forme et ses paramètres. Une bonne connaissance de la théorie probabiliste est souvent nécessaire. Les calculs exigés pour combiner plusieurs lois de probabilités sont impossibles à réaliser manuellement ou à l'aide des tableurs classiques. Les analyses de Monte Carlo permettent de résoudre le problème.

A l'occasion de leur mise en œuvre, la distinction entre variabilité et incertitude retrouve toute son importance. Elle justifie notamment la séparation dans les analyses de Monte Carlo de deux types de simulation : les simulations sur paramètres variables (nommées modélisations de première ordre) et les simulations sur paramètres incertains (qualifiées de modélisation de second ordre) La simulation de premier ordre consiste à tirer au hasard le devenir d'un patient dans l'arborescence et à répéter l'opération sur un grand nombre de patients. C'est bien la variabilité des deve-

nirs individuels qui est décrite. Dans les modélisations du second ordre la modélisation de la variabilité est imbriquée dans celle de l'incertitude. Une distribution bien définie est attribuée à chacune des variables de l'arbre. A chaque tirage les variables ont une valeur unique et différente du précédent tirage. Les simulations du second ordre sont donc les seules qui prennent réellement en compte l'incertitude. Ce sont celles qu'il convient d'utiliser dans le cadre des analyses probabilistes.

Les étapes opérationnelles de la mise en œuvre d'une modélisation de second ordre peuvent être résumées de la façon suivante.

- 1) Coder les incertitudes portant sur des grandeurs incertaines par des distributions de probabilité *a priori* qui décrivent leur fréquence d'apparition anticipée.
- 2) Pour une famille arrêtée de courbe, caractériser la valeur de ses paramètres théoriques non observables qui simule le mieux la réalité observée.
- 3) Tirer au sort chaque réalisation de la variable.
- 4) Évaluer le modèle avec les variables tirées au sort.
- 5) La moyenne des sorties du modèle approche leur espérance lorsque le nombre de tirages est grand. En réalisant B échantillons sur les couples coût ; années de vie sauvées sur la base de ces principes, il est possible de construire B rapports coût/efficacité différentiel qui peuvent être représentés dans le graphique classique à quatre cadrans.

Plusieurs types de modélisation sont envisageables. Tout d'abord, la modélisation peut, soit intégrer les résultats des mélanges de comparaisons de traitement dans un modèle de type Monte Carlo réalisé sous TreeAge®, soit être réalisée directement au décours de l'analyse statistique, dans un contexte entièrement bayésien sous WinBUGS®. La première approche, dite à deux niveaux, est la plus courante. Elle consiste à transférer les probabilités *a posteriori* dans un modèle de décision au décours de l'analyse statistique. La seconde repose sur l'emploi de modèles de décision complets intégrant à la fois le traitement statistique des données, l'analyse



des actions possibles et la quantification de leurs conséquences réalisés au sein d'un seul et même modèle bayésien (22, 23, 24, 25).

Conclusion

Exiger des moyens sur la base de données épidémiologiques ou de faits scientifiquement démontrés est indispensable mais ne suffit plus. De nouveaux critères de jugement sont à prendre en compte : le surcroît d'investissement requis pour obtenir le service attendu ou rendu ; les bénéfices supplémentaires pour le patient des dispositifs mis en place, appréhendés en condition normale d'usage et sur coupe longitudinale ; la part d'incertitude inhérente à toute décision médicale. L'analyse économique est au cœur du dialogue entre une administration qui travaille et des sociétés savantes désormais en concurrence les unes avec les autres. Une nouvelle conception de nos devoirs se forge: « DONNER À CHACUN LE MEILLEUR PAR EURO SOCIAL INVESTI ». ■

Références

- 1 - Cook DJ, Mulrow CD, Haynes RB. *Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions*. *Ann Intern Med* 1997 ; 126(5) : 376-380.
- 2 - *Cochrane handbook for systematic reviews of interventions 4.2.5*. The Cochrane Collaboration 2005
- 3 - Grimshaw J, McAuley LM, Bero LA, Grilli R, Oxman AD, Ramsay C, et al. « Systematic reviews of the effectiveness of quality improvement strategies and programmes » *Qual Saf Health Care* 2003 ; 12 : 298-303.
- 4 - Institut For Work and Health « Systematic reviews help users keep up with expanding volume of research evidence » *J Can Chiropr Assoc* 2005 ; 49(1) : 56-61.
- 5 - Moynihan R. « Evaluating health services: a reporter covers the science of research synthesis » *Milbank Memorial Fund Report* 2004.
- 6 - National Institute for Health and Clinical Excellence. *Report to the Methods Review Working Party. « Key issues arising from workshop on evidence synthesis (indirect and mixed treatment comparisons) »*. Update of the Guide to the Methods of Technology Appraisal 2007. Website : www.nice.org.uk
- 7 - Sutton AJ, Abrams KR, Jones DR « An illustrated guide to the methods of meta-analysis » *Journal of Evaluation in Clinical Practice* 2001 ; 7 (2) : 135-148
- 8 - Sutton AJ, Cooper NJ, Lambert PC, Jones DR, Abrams KR, Sweeting MJ « Meta-analysis of rare and adverse event data », *Expert Rev. Pharmacoeconomics Outcomes Res.* 2002 ; 2(4) : 367-379.
- 9 - Lu G., Ades A.E. « Combination of direct and indirect evidence in mixed treatment comparisons » *Statistics in Medicine* 2004 ; 23 (20) : 3105-3124.
- 10 - Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM « Meta-analysis of mixed treatment comparisons at multiple follow-up times ». *Stat Med.* 2007 ; 10 ; 26(20) : 3681-99.
- 11 - Boissel JP, Cucherat M « Médicament et bénéfice thérapeutique », *Dossier Médicament et Santé Publique. Actualité et Dossier en Santé Publique* 1999 ; 43
- 12 - Boissel JP « Maîtriser les indices d'efficacité d'une thérapeutique - Première partie : l'efficacité thérapeutique est une quantité », *Médecine* 2005 ; 41-45.
- 13 - Boissel JP « Maîtriser les indices d'efficacité d'une thérapeutique - Deuxième partie : apprendre à jouer avec les différentes expressions de l'efficacité thérapeutique ou d'un indice à l'autre », *Médecine* 2005 ; 86-89
- 14 - Launois R (REES-France), Béjot Y (CHU Dijon), Bourdillon F (Société Française de Santé Publique, CHU Pitié-Salpêtrière), Eckert L ((REES-France)), Fantino B (AGORA-ADIM), Finkel S (Fédération Hospitalière de France), Francesconi C (REES-France), Garassus P (Bureau de l'assurance qualité et de l'information médico-économique de l'hospitalisation privée : BAQIMEHP), Garrigues B (Collège National d'Information Médicale, Fédération Hospitalière de France et CH du Pays d'Aix), Gaubert Y (Fédération Hospitalière de France), Mégerlin F (UNIV. PARIS V), Samson Y (Service des Urgences cérébrovasculaires CHU Pitié-Salpêtrière) in *Rapport Office Parlementaire d'Évaluation des Politiques de Santé (OPEPS) septembre 2007. « La prise en charge précoce des personnes victimes d'un Accident Vasculaire Cérébral » Assemblée Nationale n° 236 - Sénat n° 475 - 27 septembre 2007 : pp 71-315*
- 15 - Sutton AJ, Abrams KR, Jones DR. *An illustrated guide to the methods of meta-analysis*. *Journal of Evaluation in Clinical Practice* 2000; 7,2: 135-148
- 16 - Launois R., Payet S., Riou-França L., le lay K., Francesconi C., Devaux M., Quémener A., Lootvoet E. () « L'évaluation des technologies de santé : les protocoles de deuxième génération » *Journal d'Economie Médicale* 2006, Vol. 24, n° 5, p. 213-228
- 17 - Bucher HC, Guyatt GH et al. « The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of clinical epidemiology* 1997 ; 50(6): 683-691.
- 18 - Hanna N, Shepherd FA, Fossella FV et al. « Randomised phase III trial of pemetrexed versus docetaxel in patients with non-small cell lung cancer previously treated with chemotherapy ». *J Clin Oncol* 2004 ; 22 : 1589-1597.
- 19 - Schuette W, Nagel S., Blankenburg T. et al. « Phase III study of second-line chemotherapy for advanced non-small cell lung cancer with weekly compared with 3-weekly docetaxel ». *J Clin Oncol* 2005 Nov 20 ; 23(33) : 8389-55.
- 20 - Camps C., Massutti B., Jimenez A. et al. « Randomized phase III study of 3-weekly versus weekly docetaxel in pretreated advanced non-small cell lung cancer a Spanish Lung Cancer Group ». *Trial Ann. Oncol.* 2005.
- 21 - Gridelli C, Gallo C, Di Maio M et al. « A randomised clinical trial of two docetaxel regimens (weekly vs 3 week) in the second line treatment of non-small cell lung cancer : the DISTAL 01 study ». *Br J Cancer* 2004 ; 91: 1996-2004.
- 22 - Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. *Bayesian methods in health technology assessment: a review*. *Health Technology Assessment* 2000; Vol 4: No 381
- 23 - Spiegelhalter DJ, Best NG. *Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling*. 2002
- 24 - Parmigiani G. *Measuring uncertainty in complex decision analysis models*. *Statistical Methods in Medical Research* 2002; 11: 513-537
- 25 - Cooper NJ, Abrams KR, Sutton AJ, Turner D, Lambert PC. *A Bayesian approach to Markov modeling in cost-effectiveness analyses: application to taxane use in advanced breast cancer*. *J.R. Statist. Soc. A.* 2003; 166, part 3:389-405