Journal d'Économie Médicale 1999, T. 17, N° 5, 343-361

L'apport indispensable de l'épidémiologie clinique aux modèles de Markov

R. LAUNOIS[1], J.-J. CROUTSCHE[2],
A. C. MÉGNIGBÊTO[1], K. Le LAY[1], C. PORTAFAX[1]

1. IREME - Institut de recherche et d'évaluation médicale et économique, Université de Paris-Nord, 74, rue Marcel Cachin, 93017 Bobigny, France.

2. CRIEGE - Centre de recherche interdisciplinaire en gestion, Université de Paris-Nord, avenue J.-B. Clément, 93430 Villetaneuse, France.

INTRODUCTION

Beck et Pauker [1] ont été les premiers à introduire les modèles de Markov dans l'analyse clinique en 1983. Cette technique permet de simuler le devenir d'un malade au décours d'une décision thérapeutique et de formuler un pronostic sur le résultat à long terme du traitement. Une modélisation simple des événements récurrents devient possible, là où les arborescences des arbres de décision classiques se densifiaient excessivement, lorsqu'elles étaient mises bout à bout pour représenter l'enchaînement des situations. Le domaine médical fournit de multiples exemples de ce type : risques de récidive dans les accidents vasculaires cérébraux [2], probabilités de rechute chez des patients schizophrènes [3], probabilités de survenue de toxicités liées à l'administration périodique de chimiothérapie dans le traitement du cancer [4-6].

Le développement de logiciels d'analyse de la décision [7, 8] permet aujourd'hui de traiter avec rapidité une masse complexe d'informations. Encore faut-il pour les maîtriser connaître les principes théoriques sur lesquels ils s'appuient ainsi que les procédures empiriques nécessaires à l'exploitation des bases de données disponibles.

La modélisation serait sans contenu et sans objet si elle ne s'appuyait pas sur la réalité clinique de l'histoire de la maladie. Le chercheur, qu'il soit économiste ou clinicien, se doit d'alimenter la simulation qu'il a construite par des données précises, exactes et validées. Un modèle détaché de l'épidémiologie clinique ne peut fournir aucun résultat fiable.

L'objectif de cette note méthodologique est triple : rappeler d'abord les fondements des processus markoviens, utiliser par ailleurs les définitions d'incidence, d'incidence cumulative et de taux

d'incidence pour expliquer les modalités de calcul des probabilités de transition, fournir enfin aux utilisateurs potentiels de ce type d'analyse, une boîte à outils opérationnelle susceptible de les aider.

MÉTHODES

1. Cadre analytique et structure des modèles de Markov

1.1. Choix des outils de recherche : Utilisation des processus stochastiques et markoviens

1.1.1. Définition des processus stochastiques

L'histoire de la maladie peut être assimilée à un processus stochastique. Son évolution par nature aléatoire s'inscrit nécessairement dans la durée. On peut la représenter à l'aide de la fonction X (t, w), t représente sa dépendance par rapport au temps, w désigne le sous-ensemble des résultats possibles qui permet de définir « l'événement » auquel s'intéresse l'observateur.

Si t est fixé au moment t = t0, X (t0, w) est un événement aléatoire dont on ne connaît pas la fréquence avec certitude mais dont toutes les possibilités de survenue ont été identifiées.

Inversement, lorsque la fonction de répartition de w est fixée, X (t, w) correspond à l'incidence cumulative de l'événement aléatoire sur la période de suivi. On parle alors de la réalisation dans le temps du processus.

Le domaine de la santé offre de multiples illustrations de ce principe. Dans un intervalle de temps prédéfini, la rechute est un événement aléatoire doté d'une probabilité de survenue donnée. Par contre, au cours du temps, la probabilité cumulative de rechute est fonction de la durée. Cette définition générale des processus stochastiques permet de différencier les processus stationnaires des processus de Markov.

1.1.2. Processus stationnaires

On dit qu'un processus est stationnaire lorsque les causes des variations observées sont purement aléatoires et indépendantes du temps.

Les pics de température relevés chez un individu s'expliquent presque toujours par l'irruption d'un agent causal exogène. Comme l'illustre la *figure 1*, leur apparition est totalement indépendante du temps.

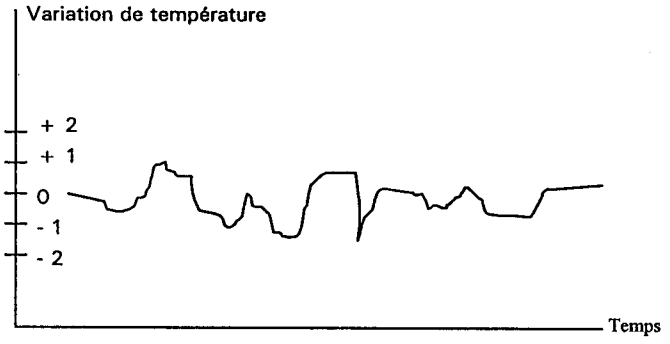


Figure 1 : Variation de la température corporelle d'un individu donné

1.1.3. Processus de Markov

Dans les processus de Markov, le temps est intégré puisque les phénomènes observés se reproduisent de façon aléatoire au fur et à mesure que celui-ci s'écoule. Il n'existe pas de lien entre les valeurs qu'une variable peut prendre à différents moments du temps. Le modèle markovien est un processus stochastique sans mémoire, le déroulement du processus est indépendant des étapes antérieures qui l'ont caractérisé.

Si les différentes fonctions de répartition des variables aléatoires à prendre en compte X (t0 ; w), X (t1 ; w)...X (tj ; w), aux instants $t0 < t1 < t2 ... t_{j-1} < tj$ sont connues, la fonction de répartition de la variable X (tj ; w), à l'instant $tj > t_{j-1}$ peut être calculée à partir de la fonction de répartition de la variable X (t_{j-1} ; w). On peut illustrer cette démarche à l'aide d'un exemple simple :

Imaginons une retenue d'eau, un barrage sur une rivière. L'eau alimente de façon régulière le réservoir de la retenue d'eau. De la même façon, l'eau s'écoule de façon régulière par les vannes du barrage.

Que se passe-t-il dans un intervalle de temps déterminé (t_{j-1}, tj) ?

- Soit X (t_{j-1} ; w), la quantité d'eau retenue par le barrage à l'instant t_{j-1},
- Soit Z (t_{j-1}, tj ; w), le volume d'eau arrivant dans le réservoir durant la période (t_{j-1}, tj),
- Soit E (t_{j-1}, tj ; w), le volume constant s'écoulant par les vannes du barrage durant le même intervalle de temps (t_{j-1}, tj).

Il est possible de calculer le volume d'eau retenu à l'instant tj :

$$X(tj) = X(t_{j-1} ; w) + Z(t_{j-1}, tj ; w) - E(t_{j-1}, tj ; w)$$

Ainsi le volume d'eau qui existe à un moment donné dans le réservoir est bien une variable aléatoire mais son niveau est entièrement déterminé par ce qui se passe dans l'intervalle de temps considéré indépendamment de tout ce qui s'est passé auparavant.

Partant de cette logique, on peut décomposer l'évolution clinique d'une pathologie en une suite d'états de santé X (t1 ; w1), X (t2 ; w2) ... X (tj ; wj) relevés à intervalles fixes pendant une durée déterminée. Parmi tous les états de santé possibles qu'un malade peut traverser, il convient de ne retenir que ceux qui correspondent à des catégories cliniques clairement définies en les croisant éventuellement avec des lieux de prises en charge ou des niveaux de dépendance. L'unité de temps adoptée pour la simulation, est fonction de la nature du problème médical étudié. Lorsqu'il s'agit d'une prise en charge au long cours, le choix de l'année semble s'imposer. Dans les maladies aiguës où les phénomènes se reproduisent sur des périodes beaucoup plus courtes, on peut adopter une période de référence plus brève. Pour les chimiothérapies, par exemple, les toxicités aiguës apparaissent immédiatement après l'administration du traitement et se résorbent dans le mois qui suit. Un découpage du temps selon le rythme d'administration des cures paraît tout à fait légitime. Quelle que soit la durée choisie comme unité de référence, l'opération a toujours pour résultat de découper la période de simulation en intervalles fixes. Ces intervalles de temps égaux sont appelés des cycles de Markov. Tous les événements susceptibles de modifier l'évolution de la maladie au cours d'un cycle sont quantifiés sous la forme de probabilités de transition. De la même manière que le niveau d'eau d'un barrage au cours d'un cycle dépend des flux d'entrée et de sortie, les chances d'être dans un état de santé spécifique sur un laps de temps donné sont déterminées par le produit des probabilités de survenue des événements qui y conduisent ou qui permettent d'y échapper.

1.2. Markov à arborescence cyclique

1.2.1. Architecture générale

L'ingénieuse présentation de Hollender [7] permet de maîtriser la complexité des cheminements en suivant leurs traces sur l'arbre de probabilités qui peut être attaché à chaque état de Markov. Les modèles de Markov présentés sous cette forme sont habituellement désignés sous le nom de **processus de Markov à arborescence cyclique**.

Quatre éléments les caractérisent :

1) L'ensemble des branches qui se greffent directement sur le nœud de Markov symbolise les états de santé mutuellement exclusifs dans lesquels le malade peut se trouver.

2) A chaque état de santé est attachée une arborescence dont les ramifications reflètent la survenue d'événements évolutifs qui vont infléchir l'évolution de la maladie dans un sens favorable ou fâcheux. Cette arborescence peut comporter soit des nœuds booléens, soit des nœuds de chance.

3) Les premiers jouent un rôle d'aiguillage. Ils commandent l'accès aux deux branches qui en sont issues et déterminent la voie qui devra être empruntée en fonction du numéro du cycle et critère de bifurcation retenu. Les seconds reproduisent le déroulement chronologique des événements intercurrents qui peuvent se produire sur le parcours indiqué.

4) Les ramifications de l'arborescence aboutissent à des nœuds terminaux. Le libellé de ces nœuds indique l'état de santé dans lequel on trouvera le malade au début du cycle suivant et sa valeur est égale au produit des probabilités relevées sur le chemin qui y mène (procédure de chaînage avant).

Les différences de coût et d'utilité qu'entraîne le passage d'un état de santé à un autre sont explicitement rattachées à chacun des états de santé qui définissent la trajectoire du malade sous traitement. L'utilité totale par cycle traversé est égale à la somme des utilités attachées à chaque état multipliée par la probabilité de s'y trouver. L'espérance totale d'utilité liée au traitement résulte du cumul des utilités enregistrées à chaque cycle sur l'horizon de vie qui a été choisi en fonction du critère d'arrêt retenu (nombre maximum de cycles ou effectif de survivants inférieurs à 1 % de la population initiale).

1.2.2. Construction graphique

La construction du modèle passe nécessairement par une étape graphique. Réunir sur un même arbre de décision, l'ensemble des éléments qui le définissent, soulève de nombreux problèmes techniques. Quel est le nombre d'états de santé à retenir ? Comment arrêter les critères qui permettent de le définir ? Quels sont la nature, le rythme, et la chronologie des événements morbides ? Quels sont les étapes et les « stop over » des parcours en boucle. Autant de questions qu'une présentation visuelle aide puissamment à résoudre : deux outils peuvent être utilisés, l'arborescence cyclique *(figure 2)* et le diagramme des états de Markov *(figure 3)*. Les schémas obtenus par la première méthode illustrent de façon ramassée comment le problème est traité par les logiciels de décision. Leur présentation dans une publication peut cependant consommer beaucoup d'espace, lorsque les événements cliniques qui expliquent le passage d'un état de santé à un autre ne sont pas simultanés dans une même période et répétitifs dans le temps, mais successifs et nouveaux. L'emploi de la technique des diagrammes de Markov peut alors se révéler utile puisqu'elle permet d'appréhender de manière panoramique la suite de la totalité des cycles thérapeutiques. Mais l'outil est beaucoup plus grossier. Aucun événement clinique n'y figure. Le maintien dans un état de santé donné est simplement indiqué par

une flèche en demi-boucle sans qu'on en connaisse la durée.

Le modèle DOPAVI [4-6] permet de montrer l'intérêt de ces deux types de présentation. L'objectif de l'étude était de confronter trois molécules : le DOcetacel, le PActitaxel, et la VInorelbine administrés à des femmes atteintes de cancer du sein en situation d'échappement thérapeutique.

19 états de santé ont été définis en fonction de la nature des réponses observées puis confirmées, du type de toxicités (aiguës ou cumulatives), et de leur degré de sévérité : toxicités mineures de grade OMS 0,1,2 et toxicités majeures de grade 3,4. A chacun de ces paramètres correspond un état de santé spécifique excepté dans le cas des toxicités cumulatives qui interdisent la poursuite du traitement quelle que soit la nature de la réponse obtenue. Les toxicités cumulatives observées avec ou sans traitement ont donc été combinées avec le type de réponses obtenues : réponse objective ou stabilité pour définir quatre états de santé, additionnels dont la pertinence sur le plan clinique semble solidement établie.

La chronologie des événements s'échelonne dans le temps en fonction du rythme des cures, du calendrier des bilans d'extension tumorale et du délai d'apparition des toxicités cumulatives. Le traitement est administré tous les 21 jours, l'intervalle de temps choisi pour la simulation est donc de trois semaines. L'identification de la réponse est faite à la fin du deuxième cycle à partir d'un bilan d'extension tumorale qui amène à un classement des malades en répondeurs et en non-répondeurs. Le bilan est répété à plus de quatre semaines conformément aux standards de l'OMS, ce qui correspond dans ce cas à la fin du quatrième cycle. Les toxicités cumulatives apparaissent après trois ou quatre cures. Il s'agit, pour le Docetaxel, d'œdème dont la durée médiane de résorption est comprise entre 18 et 26 semaines. Le modèle a donc été construit de telle sorte qu'il puisse prendre en compte ce phénomène de rémanence des toxicités dans le temps. Les toxicités aiguës ont été recensées dans les deux premiers cycles ce qui évite de multiplier le nombre des états de santé possibles tout en respectant les probabilités cumulatives de survenue de ces événements publiées dans la littérature internationale. L'emploi d'une telle procédure suppose que l'on puisse transformer les statistiques publiées à une date d'observation déterminée en un produit de probabilités de valeur équivalente calculées sur des périodes de suivi plus courtes (voir § 2).

La succession des événements dans le temps est représentée par une suite d'arbres de décision qui définissent cycle après cycle les modalités de passage d'un état de santé à un ou plusieurs autre(s) en fonction des événements cliniques qui affectent le cours de la maladie. *La figure 2* illustre les modalités d'emboîtage de ces arborescences dans le cas des toxicités aiguës. L'analyse ne porte donc que sur les deux premiers cycles avant que les réponses n'aient été identifiées.

L'arbre débute par un nœud de décision *(fig. 2-A)*. Les branches qui en émanent représentent les options thérapeutiques mises en compétition. Le nœud de Markov, qui est représenté à droite de la parenthèse par un rectangle contenant deux cercles reliés par une flèche, signale l'usage d'un processus stochastique. Les branches qui s'y rattachent correspondent à la situation de départ et aux toxicités possibles : toxicités mineures, neutropénie fébrile avec et sans hospitalisation, réactions cutanées. La probabilité d'être en situation d'échec est par définition égale à 1 sachant que toutes les femmes sont susceptibles de bénéficier d'une deuxième toxicité de traitement.

Les quatre branches suivantes du modèle DOPAVI sont activées dès le premier

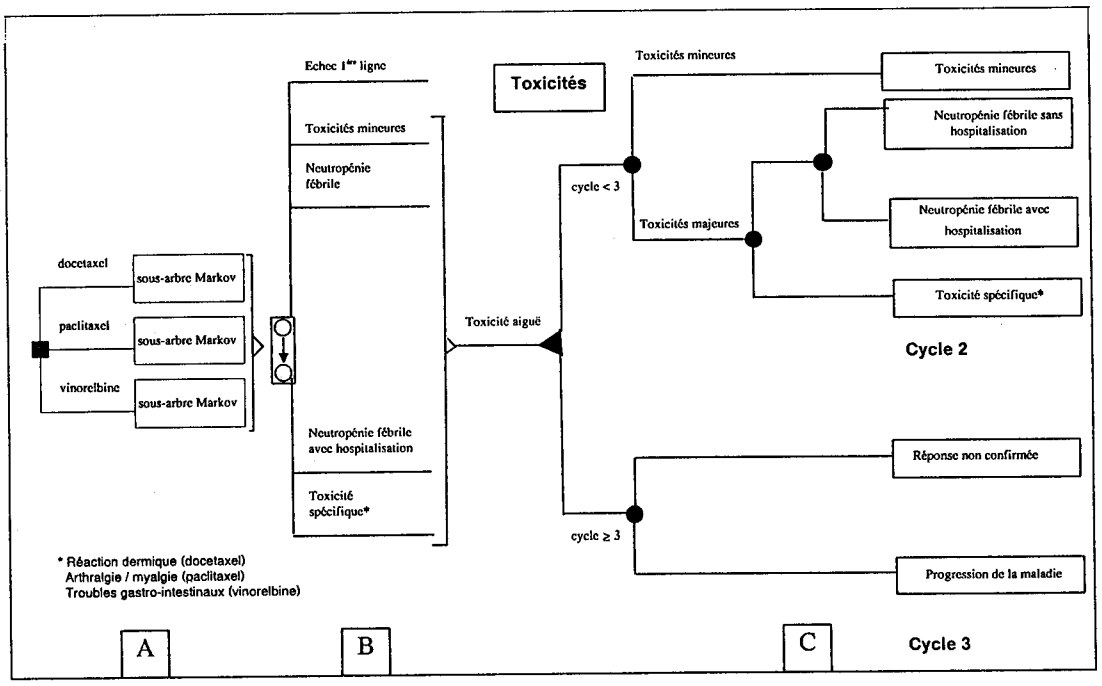


Figure 2 : Modèle de Markov dans le cancer du sein métastasé du début du traitement jusqu'au troisième cycle

cycle *(fig. 2-B)* du fait que la femme dès la première cure, peut être soumise aux événements indésirables. Les sous-groupes de population vont se scinder au premier cycle selon les risques respectifs de survenue de toxicités. Ces sous-groupes vont aboutir, à travers la combinaison des risques encourus au cours des cycles suivants, à des situations cliniques identiques ou différentes de leur situation de départ.

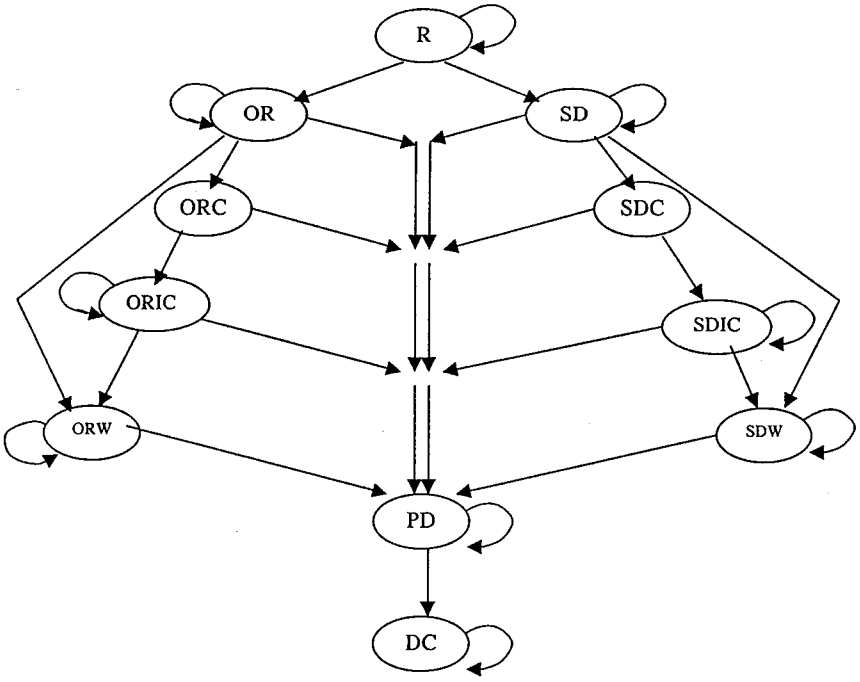
L'accolade qui figure devant l'arborescence « Toxicité aiguë » signifie que celle-ci se greffe sur chacun des états qu'elle contient. L'entrée dans l'arborescence et le cheminement sur les branches sont organisés par un nœud booléen (symbolisé par le triangle ∇). Celui-ci joue le rôle d'un aiguillage. Il précise quelle voie doit être empruntée pour passer d'un cycle au cycle suivant : la voie haute ou la voie basse. Le critère de bifurcation est donné par le nombre de simulations déjà effectuées. Si ce nombre est inférieur à 3, la sortie de l'arborescence se fait en empruntant la branche supérieure. Si ce nombre est supérieur ou égal à 3, la sortie se fait par la branche du bas.

La branche supérieure de l'arborescence porte autant de ramifications qu'il y a d'événements possibles dans l'intervalle de temps étudié. Si la patiente n'est pas réfractaire, la chimiothérapie est poursuivie. Les 4 toxicités qui étaient susceptibles de survenir au cycle 1 peuvent se reproduire au cours du second cycle. Les patientes qui n'avaient eu que des toxicités mineures au cours du premier cycle peuvent continuer à supporter leur traitement au cours du second. Mais l'inverse peut également se produire. Des

toxicités sévères peuvent apparaître chez des malades qui n'en avaient pas eu précédemment. Pour les patients ayant eu une toxicité majeure au cours du premier cycle, les mêmes effets indésirables sévères peuvent se produire une deuxième fois. D'autres malades échapperont à la toxicité sévère dont ils étaient victimes au cours du premier cycle pour retomber dans une autre toxicité sévère de même gravité, sinon de même nature. Les effets indésirables graves peuvent enfin céder la place à des toxicités mineures. Le recrutement au début du second cycle de chacun des états s'opère donc par quatre voies différentes. La fraction de la cohorte qui s'y trouve est déterminée par la somme du produit des probabilités relevées sur les chemins qui y mènent.

La sortie de l'arborescence, à la fin du second cycle, se fait après avoir enregistré les toxicités en empruntant la branche du bas vers les états de réponse non confirmée (R) ou vers la progression de la maladie (PD) qui ont été identifiées après le premier bilan.

La suite du processus est représentée à l'aide d'un diagramme de Markov *(figure 3)*. Les flèches en demi-boucle symbolisent le maintien dans le même état de santé sur plusieurs cycles. Les autres représentent le passage d'un état à un autre.



R : répondeur non confirmé ou stabilité OR : réponse objective ; SD : stabilité ; ORC : réponse objective avec toxicité cumulative ; ORIC : réponse objective avec interruption du traitement et rémanence des toxicités cumulatives ; ORW : réponse objective sans traitement et sans toxicité ; SDC : maladie stable avec toxicité cumulative ; SDIC : maladie stable avec interruption du traitement et rémanence des toxicités cumulatives ; SDW : maladie stable sans traitement et sans toxicité ; PD : progression.

Figure 3 : Schéma explicatif simplifié en arêtes de poisson

Les malades, qui ont été identifiés dans le graphique précédent comme répondeurs non confirmés ou patients stables (R) poursuivent leur traitement au début du troisième cycle. Au quatrième cycle, la chimiothérapie continue d'être appliquée aux répondeurs du cycle précédent et aux patients dont l'état reste stable. A la fin du quatrième cycle, après le bilan de confirmation, les malades sont déclarés objectivement répondeurs (OR) ou stables (SD).

- Si aucune toxicité cumulative sévère ne survient pendant le traitement, l'intégralité des 6 cures peut être administrée. Les états (OR) et (SD) se maintiennent jusqu'au terme normal de la chimiothérapie ou se vident en progression (PD). Après la fin de la chimiothérapie, la réponse ou la stabilité peut se maintenir sans traitement et sans toxicité (ORW) et (SDW) ou évoluer vers la progression (PD).
- Une toxicité cumulative de grade 3-4 peut être observée au cours du 5e cycle chez des femmes déclarées « objectivement répondeurs » ou « stables » à l'issue du bilan de confirmation de la fin du 4e cycle : (ORC) et (SDC). Le traitement est alors interrompu au début du 6e cycle, mais la toxicité persiste à faire sentir ses effets cumulatifs aussi bien chez les répondeurs confirmés (ORIC) que chez les patients stables (SDIC), puisque la durée médiane de résorption des œdèmes est comprise entre 18 et 25 semaines, soit 6 ou 8 cycles de Markov. Les mêmes phénomènes se reproduisent à l'identique jusqu'au 12e ou 14e cycle. Les autres malades passent en progression (PD).

On a supposé que le traitement n'était jamais repris après qu'il eut été arrêté pour cause de toxicité cumulative. Au fur et à mesure que les œdèmes se résorbent, les sujets peuvent continuer à répondre ou rester stables sans traitement, ni toxicité, au cours des cycles suivants : (ORW), (SDW), ou sont amenés à progresser sans toxicité (PD).

- Si la toxicité apparaît au 6e cycle, le traitement est mené à son terme chez les malades répondeurs confirmés ou stables. La 6e injection est faite. Les états (ORC) et (SDC) se débranchent au début du 7e cycle, soit en (ORIC) et (SDIC), soit en (PD) et les mêmes phénomènes que précédemment se reproduisent à l'identique.

A la fin de la simulation, les temps passés dans chacun des états traversés peuvent être cumulés et mesurés en nombre de cycles de 21 jours. Le produit du nombre cumulé de cycle-patient ajusté ou non sur la qualité de vie par la durée du cycle mesurée en fraction d'année (21/365) donne soit l'espérance de vie ajustée en fonction du confort de vie, soit l'espérance de vie normale d'un patient traité pour cancer métastatique.

1.3. Corrections dites du demi-cycle

Le modèle enregistre au cours du temps les différents états de santé qui caractérisent l'évolution de la maladie sous traitement. Intuitivement, pour mieux comprendre la nature du processus, on peut l'assimiler à une course de 400 mètres autour d'un stade.

Chaque sprinter prend place dans un couloir au moment du départ. Par hypothèse, on suppose que celui-ci est donné en ligne. Dès le début de la course, tous les concurrents s'efforcent de se glisser à la corde. Certains y parviennent immédiatement, d'autres beaucoup plus tard après avoir quasiment effectué un tour de stade. Le temps chrono est relevé pour tous au même moment, lorsqu'ils franchissent la ligne pour la seconde fois.

La performance réalisée, mesurée en distance parcourue par unité de temps, est

surestimée dans le premier cas (les coureurs réussissent à atteindre la corde immédiatement après le départ) et sous-estimée dans le second.

Pour éliminer ces biais, on peut admettre, à titre d'hypothèse, qu'en moyenne, les coureurs changent de couloir *à mi-parcours ou en milieu de cycle* [9] et qu'ils y restent pour effectuer un tour complet. Chaque état de santé peut être assimilé à un des couloirs évoqués, et la probabilité de passage d'un état de santé à un autre est, dans le modèle de Markov, évaluée en milieu de cycle, en appliquant la même règle.

Cette convention améliore la précision de la mesure. Mais, pour réussir à la mettre en œuvre, on a besoin de connaître le nombre de concurrents qui sont encore en course à mi-parcours. Tâche impossible aussi longtemps que l'on attend le passage de ceux-ci devant la tribune officielle pour en opérer le décompte. On admet donc, moyennant une seconde convention, que le nombre des « survivants » à mi-parcours est égal à celui qui a été relevé à la fin du premier tour. Simultanément, on choisit de comptabiliser le nombre de tours effectués par rapport à cette nouvelle position de référence. Tout se passe comme si, le départ étant donné à mi-parcours, les concurrents devaient effectuer un tour complet dans le couloir qui leur a été affecté avant d'avoir la possibilité d'en changer. Un tel artifice ampute la longueur de la course d'un demi-tour au départ et l'allonge d'une égale distance en fin d'épreuve.

Pour que la compétition réponde aux normes de la discipline, on impose aux concurrents de faire un demi-tour d'échauffement avant de prendre le départ et on neutralise le dernier demi-tour qui a été effectué en trop. Cette addition et cette soustraction définissent les corrections dites du demi-cycle. Sur le plan pratique, il suffit de multiplier la première ligne et la dernière ligne des traces de la simulation par 0,5, d'additionner le résultat de la première opération et de soustraire celui de la seconde du total des sommes qui apparaissent en pied de colonne de la simulation pour retrouver les valeurs données sur la façade du logiciel Tree Age [8] dans le cadre de l'analyse coût-efficacité.

2. Analyse des données médicales et calcul des probabilités de transition

Les données médicales utilisables dans un modèle se présentent le plus souvent en termes d'incidences cumulatives ou sous la forme de taux d'incidence.

Dans le premier cas, l'horizon retenu dans les essais ou les registres ne coïncide que par hasard avec la longueur du cycle choisie pour la modélisation. Il est souvent nécessaire de rétropoler les données disponibles sur des périodes plus courtes.

Dans le second cas, les taux publiés ont une valeur comprise entre 0 et l'infini, alors que les probabilités requises doivent être comprises entre 0 et 1. Il faut de ce fait transformer les taux en probabilités.

2.1. Incidence cumulative et taux d'incidence : Définitions

Au cours d'une période, un individu peut passer de l'état de non malade à celui de malade, être victime ou non d'effets indésirables, rester stable ou évoluer. Lorsqu'un individu non malade le devient ou lorsqu'une toxicité apparaît, on évoque la survenue d'un cas ou d'un événement incident. Le nombre de cas incidents observés dans une période définit *l'incidence* de la maladie, de la toxicité ou de la complication étudiée. Pour mieux spécifier sur le plan quantitatif, l'occurrence de ces événements, on peut rapporter leur fréquence par unité de temps soit à l'effectif initial des sujets observés soit au nombre potentiel d'individus à risque

pendant la période de suivi. Dans le premier cas, on obtient *l'incidence cumulative* et dans le second un *taux d'incidence.*

L'incidence cumulative est un nombre pur exprimé en pourcentage qui ne peut être interprété que si la durée du suivi est explicitement spécifiée : incidence cumulative à 2 ans, à 4 ans, etc. C'est de surcroît une probabilité conditionnelle ou un produit de probabilités conditionnelles dont « la mesure » ne peut être calculée que si les deux conditions sont satisfaisantes. L'effectif des sujets sains *en début de période* doit être exactement connu. Les sorties d'essai ou de cohorte sont toutes attribuables à la survenue des phénomènes dont on veut calculer l'incidence. Lorsque tel n'est pas le cas, il convient de mettre en place des procédures d'estimation de l'incidence cumulative sur séries incomplètes c'est-à-dire sur des données « censurées ».

Le taux d'incidence mesure un potentiel instantané de changement. En termes mathématiques, le taux est la dérivée première (dx / dt) de la variable X. En médecine, le changement est décrit par rapport à la taille de sa population exposée au risque. Si N0 représente le nombre de sujets exposés au risque au début de la période Δt (Δt allant de t0 à t) et Nt le nombre de sujets exposés au risque à la fin de la période Δt. Le taux d'incidence instantané de survenue de l'événement est égal à d (Nt) / dt. Le taux d'incidence définit donc la force d'évolutivité d'une maladie chez un patient ou sa vitesse de propagation dans une population puisqu'il mesure sur un très court *intervalle de temps,* un changement d'état de santé par rapport à un potentiel de risque.

2.2. Méthodes de calcul de l'incidence cumulative

Deux cas de figures doivent être distingués : soit l'investigateur calcule l'incidence cumulative d'emblée sur l'ensemble de la période, soit il partage celle-ci en intervalles de temps et répète ses mesures d'observations à des moments différents du temps.

2.2.1. Calcul ponctuel sur un intervalle de temps Δt

Lorsque tous les individus sont suivis pendant la période d'observation retenue (t0, t), on peut utiliser une méthode directe pour mesurer l'incidence cumulative. En revanche lorsque l'on perd de vue des patients ou lorsque l'on fixe une date de fin d'observation qui ne permet pas à tous les phénomènes de se manifester avant la clôture de l'étude, il convient de recourir à d'autres procédures. Ces procédures relèvent de la méthode des tables de survie du nom des tableaux qui sont communément utilisés pour estimer les fonctions de survie. Deux méthodes peuvent alors être envisagées suivant la manière dont on traite les perdus de vue : la méthode actuarielle et la méthode de Kaplan Meier [10-12].

• Méthode directe

Pour une cohorte fermée suivie sur une période de temps Δt allant de t0 à tj, la probabilité de survenue d'un nouvel événement pendant la durée de l'observation P (to, t) est définie par le rapport de l'incidence sur le nombre d'individus à risque en début de période.

$$P(to,t) = \frac{(No - Nt)}{No} = \frac{I}{No}$$

No étant la population exposée au risque au début de la période de temps, to

Nt, la population exposée au risque à la fin de la période de temps, t

et I, le nombre de sujets touchés par événement pendant la période.

Même lorsqu'on la calcule sur une période unique, l'incidence cumulative doit être interprétée comme une probabilité

conditionnelle puisque sa mesure suppose que soit toujours connu l'effectif de sujets à risque en début de période N0.

• **La méthode actuarielle** [10-12]

Quand les membres de la cohorte ne peuvent être tous suivis pendant la même période, la probabilité de survenue d'un événement peut alors être estimée par la méthode actuarielle qui permet de prendre en compte les sorties d'essais et les exclus vivants. La méthode est qualifiée d'actuarielle parce qu'elle repose sur l'emploi d'intervalles de temps fixes qui par nature sont identiques à ceux employés dans les calculs d'actualisation financière. On fait l'hypothèse d'une distribution uniforme des exclusions et des événements dans chaque intervalle. Ainsi, on suppose que les sujets qui ont été perdus de vue (W) sont restés en moyenne exposés au risque de survenue de l'événement pendant la moitié de la période de temps, ce qui revient à dire que la moitié des perdus de vue ont été exposés au risque de décès pendant l'intervalle complet. En conséquence, le nombre d'individus en observation est obtenu en soustrayant du nombre de vivants au début de l'intervalle N0, la moitié du nombre de perdus de vue. La probabilité de survenue d'un événement sur l'intervalle allant de t0 à t est alors calculée par la formule :

$$P(to,t) = \frac{I}{[No-(W/2)]}$$

No = nombre de sujets exposés au risque à to,

W = nombre de sujets perdus de vue pendant la période Δt

I = nombre de sujets touchés par l'événement pendant la période.

Cette probabilité est alors estimée à la borne supérieure de l'intervalle.

• **Méthode de Kaplan Meier** [10-12]

Pour mieux apprécier le profil évolutif d'une maladie, on peut présenter les proportions de patients qui n'ont pas subi l'événement après une certaine durée (exemple : fonction de survie). La présentation la plus immédiate est celle où les délais retenus correspondent à la date de survenue des événements. Dans la méthode de Kaplan Meier, le temps d'observation est divisé en intervalles inégaux, chaque intervalle débute à l'instant où un événement survient et s'arrête avant qu'il ne se reproduise. La probabilité de « survie » ou de « survie sans événement » est égale au rapport du nombre de sujets sans événement à la fin de l'intervalle sur le nombre de sujets exposés au risque au début de celui-ci. Cette méthode rapporte les exclus vivants et les perdus de vue en début d'intervalle puisqu'on considère que ces sujets ne sont pas exposés au risque dans l'intervalle. Le nombre d'individus en observation dans l'intervalle est égal au nombre de vivants en début de période moins la totalité de l'effectif des perdus de vue (W).

$$P(to,t) = \frac{I}{No-W}$$

2.2.2. Calcul sur une série d'intervalles de temps Δi

• **Calcul exact**

Par la méthode actuarielle, il est également possible de calculer la probabilité cumulée sur plusieurs intervalles de temps Δi ($1 \leq \Delta i \leq \Delta t$) ; Δi désigne l'unité de temps adoptée dont la durée arbitrairement choisie est inférieure à la fenêtre temporelle sur laquelle la probabilité cumulée doit être calculée ; l'indice i désigne l'un quelconque de ces intervalles. Les probabilités Pi pour chaque intervalle de temps sont d'abord calculées de point en point en utilisant une des trois mé-

thodes présentées ci-dessus. Le produit des probabilités complémentaires de non-survenue de l'événement (1–Pi) une fois déduite de l'unité nous donne la probabilité cumulée de survenue de l'événement sur l'horizon retenu (t0, tj).

$$P(to,tj) = 1 - \prod_{i=1}^{i=j} (1 - Pi)$$

Pi = probabilité de survenue d'un événement sur l'unité de temps choisie.

1–Pi = probabilité de survenue d'un événement complémentaire.

Π = produit des (1–Pi).

Supposons que l'unité de temps adoptée Δi soit égale à 1 an. On calcule d'abord la probabilité de survivre dans la première, la deuxième et la troisième année. Pour survivre à 3 ans, il faut bien évidemment avoir survécu à chacun des intervalles de temps pris en compte d'où le produit des probabilités. Finalement on estime la probabilité de décéder avant 3 ans en déduisant la probabilité d'être vivant en début d'étude qui par définition est égale à 1.

$$P(0,3\ ans) = 1 - \prod_{i=1}^{3} (1 - Pi)$$

La probabilité de décès dans chacun des intervalles Δi (i = 1, 2 ou 3) s'élève respectivement à 0,2, 0,3, et 0,4. Si les individus sont vivants au moment de rentrer dans l'étude : P(0) = 1, alors la probabilité de décès avant 3 ans sera :

$$P(0,3) = 1-[(1-0,2)*(1-0,3)*(1-0,4)]$$

$$P(0,3) = 1-[0,8*0,7*0,6] = 1 - 0,336 = 0,664$$

• **Formules simplifiées**

Si on fait l'hypothèse que la probabilité pour chaque intervalle de temps Δi est constante au cours du temps pendant la période définie, la probabilité complémentaire de non-survenue de l'événement pour l'ensemble des intervalles i allant de 0 à j s'écrit $(1 - Pi)^j$. L'incidence cumulative de la date d'entrée dans l'étude à la fin de celle-ci, peut alors être calculée à partir de la formule suivante :

$$P(t0, tj) = 1 - (1 - Pi)^j$$

Cette relation permet de passer de l'incidence cumulative sur un intervalle Δi à l'incidence cumulative P(t0, tj) sur une période allant du moment t0 au moment tj, et inversement.

En conséquence, pour estimer la probabilité de survenue d'un événement sur un intervalle i, il suffit de prendre la $j^{ème}$ racine de la probabilité de non-survenue de l'événement pendant la période d'observation allant de t0 à tj, $1 - P(to, tj)^j$. On obtient alors :

$$(1 - Pi) = \sqrt[j]{(1 - P(to, tj))}$$

Les termes de cette égalité peuvent alors être reclassés pour obtenir la probabilité ponctuelle pour un intervalle de temps.

$$Pi = 1 - [1 - P(t0, t)]^{1/j}$$

j = le nombre de périodes Δi = tj–t0 / Δi.

Cette formule est très largement utilisée dans la construction des modèles de Markov. Si l'on divise la période d'observation d'un essai telle qu'elle est déterminée par sa date de début t0 et sa date de fin tj par l'unité de temps Δ qui a été choisie par le modélisateur pour définir la longueur du cycle de Markov, alors on transforme toutes les incidences cumulatives mesurées sur les différents critères de l'essai en un produit de probabilités qui a exactement la même valeur, mais dont les éléments sont mesurés sur des intervalles de temps correspondant au pas de la simulation. Ces pro-

babilités par unité de temps sont communément appelées par les économistes des probabilités de transition.

2.3. Mesure du taux d'incidence

Les mesures instantanées c'est-à-dire calculées sur un très court intervalle de temps sont très difficiles à mettre en œuvre sur la plan pratique. Le taux d'incidence est le plus souvent estimé à partir d'observations faites pendant une période déterminée sur une population donnée. Les travaux de Beck et de Pauker plus connus sous le nom de méthode DEALE [13-14] montrent toutefois que l'outil mathématique offre un moyen opérationnel d'en faire le calcul. Dans les deux cas, on parlera de *taux moyen d'incidence* sur une période allant du temps t0 au temps t, et l'on notera ce taux $\mu(t0,t)$.

2.3.1. Estimation empirique : Personnes - Années à risque

Le taux mesuré sur une cohorte entière est estimé par le rapport de l'incidence I sur le nombre cumulé de Personnes-Années à risque (PT) dans la population en observation pour la période.

$$\mu(to,t) = I/PT$$

Le dénominateur de ce rapport s'exprime donc en unités complexes qui comprennent à la fois le nombre de personnes à risque et pour chacune d'entre elles la durée du risque pour la période d'observation. L'exposition au risque commence à la date d'entrée dans l'étude et cesse le jour où survient l'événement dont on veut étudier le taux d'incidence. En portant sur le plan cartésien le nombre de personnes en ordonnée et le temps de suivi en abscisse on voit que le concept de personne-temps à risque est représenté par une surface d'où le nom de densité d'incidence qui est parfois donné au taux calculé à partir d'un tel dénominateur.

Le taux d'incidence n'est pas un nombre pur, il porte la marque des unités de mesure à partir desquelles il est évalué. Sa valeur numérique change avec l'unité de temps retenue ; par exemple 1 000 personnes années à risque a la même valeur que 100 décades à risque.

2.3.2. Estimation mathématique par la méthode DEALE

• Calcul sur un intervalle de temps

Lorsque les données disponibles sont complètes et sous réserve que le taux d'incidence μ soit constant, on peut démontrer que le nombre d'individus qui sont affectés par la survenue d'un événement à la date t est une fonction exponentielle décroissante du nombre de sujets à risque au début de la période :

$$Nt = No * \exp[-\mu(t-to)]$$

La probabilité de survenue d'un événement dans la période (t0,t) est, compte-tenu de la définition de l'incidence cumulative adoptée, égale à :

$$P(to,t) = 1 - \frac{Nt}{No}$$

Soit après substitution :

$$P(to,t) = 1 - \exp^{-\mu(t-to)}$$

μ = taux moyen estimé
$(t - t0)$ = durée de la période

Si μ symbolise par exemple le taux de décès observé du temps t0 au temps t, $\exp^{-\mu[to,t]}$ mesure le nombre de survivants et $1 - \exp^{-\mu[to,t]}$ représente la probabilité de décéder avant la fin d'étude.

• Calcul sur une série d'intervalles

Si les données sont censurées et si les taux d'incidence varient dans le temps, l'estimation des fonctions de survie se fait point par point. Sa période d'observation est alors divisée en autant d'inter-

valles qu'il y a de taux d'incidence différents. La probabilité cumulative sur l'ensemble des intervalles Δi allant de t0 à tj est évaluée à l'aide de la formule :

$$P(to,t) = 1 - \exp^{\left[-\sum_{i=1}^{i=j} \mu_i \Delta_i\right]}$$

Δi = le ième intervalle.

μi = le μ pour le ième intervalle.

Σ = somme arithmétique des (μi*Δi).

Si les taux d'incidence ne varient pas au cours d'une période donnée, distinguer les intervalles de temps au cours de cette période n'a pas d'intérêt, puisque le taux μ est le même dans chacun d'eux. Il reste toutefois à choisir l'unité temporelle de référence, j indique le nombre de ces unités contenues dans l'intervalle t0 t. La formule simplifiée s'écrit :

$$P(to,t) = 1 - \exp^{(-\mu j * j)}$$

μ = taux d'incidence constant sur l'intervalle de temps Δt.

j = nombre d'unités calendaires choisi sur l'intervalle Δt.

Lorsque la probabilité cumulative de survenue d'un événement sur une période d'observation allant de t0 à t est connue (par exemple, une probabilité cumulative de survie à 5 ans égale x pour cent – un temps médian jusqu'à progression égal à x mois), on peut calculer à l'aide de cette formule :

1) le taux d'incidence pour d'autres unités temporelles (le mois ou l'année, soit j respectivement égal à 60 et à 5),

2) la probabilité P qui lui est associée dans le cadre de l'unité de temps choisie.

On pose le calcul dès que l'on est en position de chiffrage de la probabilité cumulative de survenue de l'événement sur la période d'observation t0,t.

$$\exp^{(-\mu j * j)} = 1 - P(to,t)$$

Ce qui donne en prenant le logarithme népérien des deux termes de l'égalité :

$$\mu j = \frac{-ln(S)}{j}$$

Sachant que S = 1 – P(t0,t) mesure la probabilité complémentaire de non-réalisation de l'événement, la probabilité de survenue du risque P qui correspond à ce taux d'incidence μj sur l'unité temporelle retenue est calculée avec la formule générale :

$$P = 1\ e^{(-\mu j * 1)}$$

j étant égal à 1 puisque l'on ne raisonne que sur la seule période calendaire de référence.

L'approximation de la fonction de survie par une fonction exponentielle décroissante est un autre procédé de calcul qui permet d'aboutir au même résultat que la méthode actuarielle en employant le taux d'incidence comme variable intermédiaire.

3. Estimation de la qualité de vie

A l'évidence, en matière de chimiothérapie, on ne peut dissocier la durée de survie de sa qualité [15]. A chacun des 19 états de santé, qui caractérisent l'évolution du malade sous traitement, correspond un niveau de qualité de vie, qui est évalué sur une échelle comprise entre 0 et 1 ; 0 désignant le décès et 1 l'état de parfaite santé. Pour évaluer la qualité de vie des patients, il suffit de leur demander où ils se situent entre ces deux extrêmes, ce qui revient à déterminer la valeur d'un coefficient de qualité de vie pour chacun des états traversés. Ces coefficients de qualité de vie sont autant de facteurs d'ajustement de la quantité de vie en fonction de sa qualité. Si le coefficient est égal à 0,5, passer un cycle dans cet état est équivalent à passer

0,5 cycle dans un état de bonne santé. Plus explicitement, on dira qu'une année passée dans cet état ne vaut pas 12 mois mais 6. Lorsque l'on ne s'intéresse qu'à la survie, on considère que l'utilité de tous les états, autre que le décès, est de 1. A chaque cycle, les probabilités de transition déterminent la proportion de patients répartis entre les états de santé. Cette probabilité, multipliée par l'utilité de l'état correspondant, définit l'espérance d'utilité associée à cet état.

Chaque état clinique du modèle DOPAVI fait l'objet d'une description de son retentissement en termes de qualité de vie. L'indicateur de santé utilisé comme canevas a été le Heath Utility Index (HUI : Mark 2 [16] et Mark 3 [17]). Ce système multi-attribut de classification comporte huit dimensions, les cinq dimensions pertinentes par rapport à la maladie ont été conservées, les autres ont été éliminées. L'intensité des plaintes a été mesurée sur des échelles à cinq ou six niveaux. Les descriptifs ont été revus par un groupe français de 8 experts en fonction de leur expérience clinique (5 médecins du Centre de Lutte Contre le Cancer « Oscal Lambret » à Lille et 3 infirmiers de « Santé-Service » à Paris). Le profil obtenu définit un stéréotype d'état de santé communément associé selon les experts cancérologues à l'entité clinique considérée. Ces stéréotypes doivent être classés entre 0 et 1 en fonction du jugement subjectif que les malades ou leurs représentants portent sur les épreuves traversées. La mesure quantitative de ces jugements a été réalisée dans le cadre d'une enquête d'utilité effectuée auprès de 20 infirmières sur 3 sites, dont deux sites hospitaliers (Lille et Villejuif) et un site ambulatoire (Santé Service). Les techniques de révélation des préférences mises en œuvre étaient celles du Standard Gamble et de l'Echelle Visuelle Analogique, introduites au Canada dans les années 70 par Torrance (Mc Master University). Les coefficients de pondération ainsi dégagés ont été utilisés pour calculer une espérance de vie sans progression ajustée en fonction de la qualité de vie.

RÉSULTATS

1. Estimation de la valeur des probabilités simples ou cumulatives à partir des données publiées

1.1. Taux de réponses et de toxicités publiées

Les données dont nous disposons provenaient des résultats du dossier d'enregistrement, c'est-à-dire du pooling des essais réalisés sur 111 patientes prétraitées essentiellement sur des deuxièmes lignes avec 86 % de patientes anthracycline résistantes.

Les hypothèses initiales du modèle reposent donc sur les données suivantes *(cf. tableau I)*.

1.2. Probabilités de transition

1.2.1. Définition

Une probabilité de transition est égale à la somme du produit des probabilités relevées sur les chemins qui partent de tous les états de santé possibles pour arriver à un état de santé donné.

Le cas des toxicités dans le modèle DOPAVI donne une bonne illustration de cette procédure de calcul. L'état toxicité mineure, par exemple, peut s'auto-reproduire ou être alimenté par des passages en provenance des états neutropénie fébrile avec hospitalisation, neutropénie fébrile sans hospitalisation ou de réaction cutanée. Sachant que les probabilités au premier cycle d'être atteint par ces toxicités s'élevaient respectivement à : 0,861, 0,094, 0,0095 et 0,036 et que le risque de passer en toxicité mineure au second cycle quelle que soit la toxicité relevée au premier est de 0,861, la pro-

Tableau I : Résultats de la méta-analyse des essais docetaxel [18-20]

Variables	Valeurs publiées	Nature du paramètre	Probabilités
Efficacité			
Réponse objective	57,1 %	Proportion	0,571
Durée médiane de réponse	28 semaines	Taux DEALE	0,159*
Durée médiane de survie	12 mois	Taux DEALE	0,039
Temps jusqu'à Progression	21 semaines	Taux DEALE	0,435
Toxicités *(sur les 2 premiers cycles)*		Proportion	
Neutropénie fébrile non hospitalisée	1,9 %	Proportion	0,0095
Neutropénie fébrile hospitalisée	17,9 %	Proportion	0,094*
Réaction cutanée	7 %	Proportion	0,036
Toxicités mineures	73,2 %	Proportion	0,861
Arrêt du traitement pour œdème	1,9 %	Proportion	0,006

* Les modalités de passage sont expliquées pour ces deux valeurs, à titre d'exemple, dans les § 1.2.2 et 1.2.3.

babilité d'être au deuxième cycle en toxicité mineure est égale à :

$$(0{,}861*0{,}861) + (0{,}0095*0{,}861) + (0{,}094*0{,}861) + (0{,}036*0{,}861) = 0{,}861$$

On retrouve bien le chiffre global de toxicités mineures publié, son incidence a été éclatée sur deux cycles.

La probabilité qui s'attache à chaque état indique la proportion des malades qui s'y trouvent. Si l'on multiplie sa valeur par l'utilité correspondante, on obtient l'utilité de cet état. L'utilité de chacun des états est additionnée sur l'ensemble de ceux-ci (Utilité du cycle), puis s'ajoute à la production de santé du cycle précédent afin de calculer l'utilité totale après trois simulations.

1.2.2. Probabilité de survenue des neutropénies fébriles hospitalisées

Les probabilités concernant les toxicités des patients traités avec le docetaxel sont calculées sur l'ensemble des patients traités avec docetaxel qu'ils soient prétraités ou non, anthracycline-résistants ou non. Il y a 17,9 % de neutropénies fébriles hospitalisés au total, l'incidence cumulative sur 2 cycles est donc égale à 0,179. La probabilité correspondante pour un cycle P est égale à $1-(1-P_{0,2})^{1/2}$ soit $1-(1-0{,}179)^{1/2} = 0{,}094$.

1.2.3. Probabilité de maintien de la réponse objective

Les malades qui ont été objectivement déclarés répondeurs à la fin du quatrième cycle, peuvent rechuter dès la fin du cinquième. La durée médiane de réponse, partielle ou complète, qui a été relevée sur les données poolées des trois essais docetaxel est égale à 28 semaines, soit 9,33 cycles (28/3). Cette donnée était calculée par rapport à la date de début de traitement. Dans le modèle DOPAVI, elle a été évaluée par rapport à la date de confirmation de la réponse, soit un écart de 15 semaines, c'est-à-dire de 5 cycles entre les deux estimations. Les chiffres retenus dans le modèle ont été corrigés pour tenir compte de cette différence. La durée médiane de traitement après confirmation est égale à 4,33 (9,33–5). Le taux par cycle du maintien de la réponse est égal :

$$\mu = \frac{-\ln(\textit{Durée médiane de réponse})}{4{,}33\ \textit{cycles}}$$

Par définition la durée médiane de réponse correspond à la probabilité d'avoir 50 % des patients non rechutants.

$$\mu = \frac{-ln\,(0,5)}{4,33} = 0,160\,par\ cycle$$

Soit une probabilité de rester répondeur :

$$P = 1 - e^{-0,160} = 0,148$$

La probabilité de rechuter par cycle est alors égale à 1–P soit 0,852 par cycle.

2. Mesure de la qualité de vie dans chaque état de santé

Le classement des utilités en fonction de la réponse *(tableau II)* reflète bien le confort que celle-ci apporte lorsqu'elle est confirmée (0,81) et le poids des souffrances qu'entraînent l'absence de changement (0,74), et surtout la progression (0,65).

Les résultats mettent bien en évidence l'interaction des phénomènes de réponse et de toxicité dans l'appréciation du confort de vie du patient : les toxicités cumulatives sont mieux supportées si la réponse est positive. Ainsi, lorsqu'il y a apparition d'œdème, la qualité de vie reste relativement élevée chez le patient répondeur (0,74) ou stable (0,73), tandis que chez le patient en progression, elle s'abaisse à un score de 0,58. La survenue d'une neuropathie est fortement pénalisante lorsqu'elle est conjuguée à l'absence de changement ou à la progression du cancer (0,50) alors que cette toxicité est mieux supportée chez le patient répondeur (0,57).

3. Calcul de l'espérance de vie ajustée sur la qualité

Trois sous-ensembles d'états de santé ont été distingués selon qu'ils correspondent à des toxicités aiguës ou cumulatives (TOX), à un temps de rémission (TWiST) [21] ou à une progression (REL).

La durée de la rémission est obtenue par simple sommation du laps de temps passé dans les états de réponse sans toxicités et sans symptômes (TWiST), et de la durée pendant laquelle il a dû subir les effets aigus ou cumulatifs des toxicités (TOX).

Au terme du processus, on obtient un nombre de cycles sans progression qui s'élève respectivement à 8,22. Ces cycles-patients de 21 jours peuvent être transformés en espérance de vie sans pro-

Tableau II : Classement des utilités en fonction des états de santé

ÉTATS CLINIQUES	Stéréotypes d'états de santé	Coefficients de qualité de vie (standard gamble)
Réponse confirmée	ES5	0,81
Réponse confirmée avec œdèmes graves	ES14	0,74
Réponse confirmée et neuropathie sévère	ES8	0,57
Stabilité	ES6	0,75
Stabilité et œdèmes graves	ES15	0,73
Progression	ES7	0,65
Progression et interruption du traitement pour œdèmes graves	ES19	0,58
Progression et neuropathie sévère	ES10	0,50

Note : Pour les utilités associées aux autres états de santé voir la référence [4].

gression (disease-free survival in years) et en jours sans progression (disease-free survival in days). Dans le premier cas, il suffit de les multiplier par la fraction d'année que représente un cycle de 21 jours, et dans le second par 21 jours. On obtient 0,473 année ou 173 jours sans progression. Ces mêmes résultats ajustés en fonction de la qualité de vie font apparaître une espérance de vie sans progression et sans inconfort de 125 jours.

Tableau III : Résultats du traitement avec intégration de la qualité de vie à la survie

	Traitement docetaxel
Espérance de vie sans progression	173 jours
Espérance de vie sans progression et sans inconfort	125 jours

4. Discussion

Le plan expérimental constitué de 111 personnes ne permet pas en l'état actuel de conclure avec une absolue certitude. En effet les règles classiques d'échantillonnage s'appuient sur le principe des 5 % d'erreurs. Or en s'appuyant sur une distribution gaussienne et sur une population mère importante, l'échantillon devrait se composer de 400 personnes. Nous sommes loin de ce nombre et le risque d'erreur associé à la proportion de 0,5 (cas le plus défavorable) avec N = 111 est de 9,49 %.

Il est vrai qu'à ce stade de développement du produit, la possibilité de réunir un tel effectif est très faible. Par ailleurs le recrutement d'une telle population demanderait plus de temps pour un gain d'information relativement modeste.

De surcroît les techniques ont toutes les chances d'évoluer avant que le plan d'expérimentation soit clos.

L'avenir est en deuxième ligne aux associations d'agents cytotoxiques et les mono chimiothérapies sont à terme condamnées. En tout état de cause, les vrais débouchés sont pour les industriels en première ligne de traitement voire en traitement adjuvant. Dans un tel contexte, l'analyse de décisions offre un outil de simulation peu coûteux à l'emploi et instructif dans ses enseignements.

CONCLUSION

Le processus de Markov autorise une souplesse que ne permettent pas les chaînes de Markov classiques puisqu'il est possible d'utiliser des probabilités dont la valeur change au fur et à mesure de l'écoulement du temps.

L'affectation d'un coefficient de qualité de vie à chaque état permet d'intégrer les paramètres de qualité de vie dans la prise de décision. De façon symétrique, l'attribution de « vignettes de coût » [22] aux différents états cliniques traversés permet d'estimer le montant prévisionnel des dépenses entraînées par la prise en charge des patientes du début du traitement au décès. La confrontation de ces deux types de données débouche tout naturellement sur le calcul du ratio coût-efficacité différentiel.

Compte tenu de leur souplesse d'utilisation, les processus de Markov sont appelés à jouer un grand rôle dans l'analyse médico-économique.

BIBLIOGRAPHIE

[1] Beck RJ, Pauker SG. The Markov process in medical prognosis. *Medical Decision Making* 1983;3:419-458.

[2] Launois R, Menigbeto A, Le Lay K, Scart-Grès C, Gaudin A-F. Coût d'un suivi à 5 ans des Accidents Vasculaires Cérébraux en fonction du degré d'autonomie et des structures de prise en charge. *Acte du colloque « Econométrie de la santé XII -*

Système Nerveux Central ». Vérone, 8-9 juillet 1999.

[3] Launois R, Schulenburg M, Knapp M, Toumi M. « Cost-effectiveness of sertindole versus olanzapine or haloperidol : A comprehensive model ». *International Journal of Psychiatry in Clinical Practice* 1998;2 (Suppl 2):S79-S86.

[4] Launois R, Reboul-Marty J, Henry B, Bonneterre J. « A cost-utility analysis of second-line chemotherapy in metastatic breast cancer : docetaxel versus paclitaxel versus vinorelbine ». *PharmacoEconomics* 1996;*10*;5:504-521.

[5] Launois R, Reboul-Marty J, Henry G, Bonneterre J. « Cost utility in second line metastatic cancer ». The Authors' reply. *PharmacoEconomics* mai 1997;*11*;5:495-497

[6] Launois R, Reboul-Marty J, Bonneterre J et al. « Evaluation médico-économique de la chimiothérapie de deuxième ligne dans le cancer du sein métastatique : comparaison du docetaxel, du paclitaxel et de la vinorelbine ». *Bulletin du Cancer* 1997;*84*; 7:709-21.

[7] Hollenberg J. *SMLTREE : The all purpose decision tree builder.* Boston : Pratt Medical Group 1993, 76 p.

[8] Data ™ 3.5 for Health Care ; user manual Tree Age software Inc., 1075, Main Street, Williams Town, MA 01267.

[9] Sonnenberg FA, Beck JR. Markov models in medical decision making. A practical guide. *Medical Decision Making* 1993;13: 322-38.

[10] Laplanche A, Com-Nougue C, Flamant R. *Méthodes statistiques appliquées à la recherche clinique.* Paris, Flammarion, 1991.

[11] Hill C, Com-Nougue C, Kramar A et al. *Analyse statistique de données de survie.* Paris, Flammarion, 1990.

[12] Kleinbaum DG, Kupper LL, Mogenstern H. *Principles and quantitative methods. Epidemiologic research.* London : Lifetime Learning Publications, 1982, 529 p.

[13] Beck JR, Pauker SG, Gottlieb JE et al. A convenient approximation of life expectancy (the « DEALE »). II - Use in Medical Decision Making. *Am J Med* 1982;73:889-897.

[14] Launois R. « Analyse coût-efficacité des stratégies thrombolytiques ». *Archives des Maladies du Cœur et des Vaisseaux* 1989;82(NSIII):55-62.

[15] Launois R. Les mesures de la Qualité de Vie. In C. Jasmin, J. Lévy, G. Bez. *Cancer Sida ...La Qualité de Vie.* Les empêcheurs de tourner en rond. Ed. Synthélabo Groupe Paris. 1996:165-175.

[16] Feeny D, Furlong W, Barr RD et al. A comprehensive multi-attribute system for classifying the Health Status of Survivors of Childhood. *Cancer J Clin Oncol* 1992; *10*;6:923-928.

[17] Furlong W, Feeny D, Torrance GW et al. Guide to design and development of health-state utility instrumentation. *CHEPA Working Paper Series n° 90-9*, Ontario : McMaster University, 1990, 140 p.

[18] Ten Bokkel Huinink WW, Prove AM, Picard M et al. A phase II trial with docetaxel in second line treatment with chemotherapy for advanced breast cancer : a study of the EORTC Early Clinical Trials group. *Ann Oncol* 1994;5;6:527-32.

[19] Valero V, Holmes FA, Walters RS et al. Phase II Trial of docetaxel : a new, highly effective antineoplastic agent in the management of patients with anthracycline-resistant metastatic breast cancer. *J Clin Oncol* 1995;13:2886-94.

[20] Ravdin PM, Burris HA, Cook G et al. Phase II Trial of docetaxel in advanced anthracycline-resistant or anthracenedione-resistant metastatic breast cancer. *J Clin Oncol* 1995;13:2879-85.

[21] Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and Benefits of Adjuvant Therapy in Breast Cancer : A Quality-Adjusted Survival Analysis. *J Clin Oncol* 1989;7;1:36-44.

[22] Launois R, Portafax C, Perrocheau G. Cancer du colon - Chimiothérapie de seconde ligne : les enjeux. *Le Moniteur hospitalier* 1999;112:18-24.

ANALYSE

BULLETIN D'INFORMATIONS EN ÉCONOMIE DE LA SANTÉ - QUESTIONS D'ÉCONOMIE DE LA SANTÉ

CREDES, 1, rue Paul Cézanne, 75008 PARIS

Abonnement 1 an (environ 10 numéros) : 300 FF

ISSN : 1283-4769

Le CREDES a réalisé une étude sur la concentration des dépenses et les grands consommateurs de soins médicaux. Cette étude a pour objectif de déterminer comment se concentre la dépense médicale et quelles sont les caractéristiques des personnes qui consomment beaucoup pour leur santé.

Cette étude propose ainsi, à partir de l'analyse des dépenses médicales présentées au remboursement de la Sécurité Sociale en 1995, une analyse de la concentration d'une partie des dépenses médicales estimées à environ 8 100 francs, dont 40 % de frais hospitaliers et 60 % de soins ambulatoires.

Les consommations médicales varient fortement selon différents critères : état de santé, âge, sexe, milieu social et couverture maladie. Il apparaît ainsi qu'un quart de la population est à l'origine de 80 % des dépenses totales et que 5 % seulement des personnes en concentrent près de la moitié.

La concentration des dépenses varie selon le type de soins

La dépense de soins hospitaliers est plus concentrée que celle de soins ambulatoires puisqu'elle concerne un nombre moins important de personnes (16 %) mais la dépense engagée est très élevée, environ 21 300 francs par personne. 10 % des plus grands consommateurs de soins hospitaliers concentrent 97 % des dépenses hospitalières. En revanche, ces mêmes consommateurs ne concentrent que 43 % des dépenses ambulatoires.

En ce qui concerne le type de consommations, on note deux faits marquants : (1) on trouve des consommations très courantes (médecins, pharmacie) qui concernent plus de 80 % des personnes au cours d'une année, (2) certaines consommations ne concernent qu'un faible pourcentage de personnes (sages-femmes, optique médicale, soins de dentistes).

La composition du panier de soins se modifie avec la progression des dépenses

Quand le niveau de la consommation médicale globale s'élève, les structures des dépenses par type de soins se déforment. Pour les plus grands consommateurs, les frais hospitaliers augmentent vite pour atteindre des montants plus élevés que les frais liés aux soins de ville. Les soins de ville complète l'arsenal diagnostique et thérapeutique utilisé par les grands consommateurs.

Le profil des 10 % plus grands consommateurs

Ces personnes concentrent 70 % des remboursements de la Sécurité Sociale. En moyenne, chaque personne dépense environ 48 300 francs de soins médicaux, dont 63 % de soins hospitaliers.

Les 10 % plus grands consommateurs déclarent 1,8 fois plus de maladies que les autres (tumeurs, maladies cardiovasculaires, maladies ophtalmiques).

La proportion de grands consommateurs croît avec le risque vital et l'invalidité. La quasi-totalité des soins hospitaliers et ambulatoires augmentent avec le risque vital ou l'invalidité. Par rapport à l'ensemble de la population appariée, ces grands consommateurs comportent une majorité de femmes (59 % *versus* 50 %). Ils sont plus âgés (âge moyen 51 ans *versus* 35 ans), et plus d'un tiers d'entre-eux sont des personnes âgées de plus de 65 ans.

Parmi ces personnes, 85 % bénéficient d'une couverture sociale complémentaire, soit la même proportion que dans la population générale.

Pour en savoir plus :

- ***Concentration des dépenses et grands consommateurs de soins médicaux***
  L. Com-Ruelle, Sylvie Dumesnil
  Biblio n° 1269 – Prix : 210 FF
- ***Profils de consommateurs de soins***
  M.-J. Sourty-le-Guellec
  Biblio n° 1246 – Prix : 130 FF

Christell JULIEN