

*COLLÈGE NATIONAL  
DES GYNÉCOLOGUES ET OBSTÉTRICIEUS FRANÇAIS  
Président : Professeur J. Lansac*

# **Extrait des Mises à jour en Gynécologie Médicale**

—

**Volume 2009  
publié le 9.12.2009**



*TRENTE-TROISIÈMES JOURNÉES NATIONALES  
Paris, 2009*

# Intérêt des prédicteurs clinicopathologiques dans les cancers du sein

C. COUTANT, M. BALLESTER, E. DARAI, S. UZAN, R. ROUZIER \*  
(Paris)

## Résumé

*L'utilisation de prédicteurs clinicopathologiques permet d'optimiser la prise en charge des malades en identifiant des sous-groupes spécifiques de patientes. Ils offrent la possibilité de proposer aux patientes une information pertinente et objective mais surtout personnalisée tant sur le plan pronostique (risque de récurrence) que prédictif (réponses aux traitements adjuvants ou à la chimiothérapie néoadjuvante).*

*Des prédicteurs clinicopathologiques ont été développés dans quatre grands domaines : I) Adjuvant! Online (i.e. prédiction à 10 ans de la survie globale et de la survie sans récurrence ; prédiction du bénéfice attendu des différents traitements adjuvants) ; II) la prédiction de la réponse à la chimiothérapie néoadjuvante ; III) la prédiction du statut des ganglions non sentinelles (GNS) en cas d'envahissement du ganglion sentinelle (GS) ; et IV) la prédiction de l'envahissement d'au moins 4 ganglions axillaires en cas d'envahissement de 1 à 3 GS.*

\* Hôpital Tenon - APHP - Université Pierre et Marie Curie Paris 6 - Service de gynécologie obstétrique et médecine de la reproduction - UPRES EA 4053 - 4 rue de la Chine - 75020 Paris

Correspondance : charles.coutant@tnn.aphp.fr

*Cependant, une validation rigoureuse de ces prédicteurs est indispensable avant toute utilisation en pratique clinique. Les paramètres à prendre en compte sont la discrimination, la calibration, le concept d'utilité clinique et la comparaison à un modèle optimal. Par ailleurs, les caractéristiques démographiques et clinicopathologiques (découverte de nouveaux biomarqueurs par exemple) ainsi que les protocoles de traitements - néoadjuvants et adjuvants - évoluent rapidement. Ainsi, ces différents prédicteurs doivent être régulièrement implémentés et optimisés afin de conserver leur pertinence.*

*Les perspectives sont essentiellement représentées par la prédiction spécifique et la prédiction combinée qui apparaissent de plus en plus comme des éléments incontournables dans le domaine de la prédiction en cancérologie mammaire.*

*Mots clés : cancer du sein, prédiction, modèles prédictifs, nomogramme, score*

## INTRODUCTION

L'utilisation de prédicteurs clinicopathologiques permet d'optimiser la prise en charge des malades en identifiant des sous-groupes spécifiques de patientes. Dans les groupes à risque, les patientes pourront bénéficier d'une intensification des traitements ou du développement de thérapeutiques spécifiques avec pour but la diminution du taux de récurrence et une amélioration de la survie. Dans l'idéal, elle permettrait de proposer des traitements « à la carte ». A contrario, dans les groupes à bas risque, une désescalade thérapeutique tant sur le plan médical que chirurgical permettrait de diminuer la morbidité.

L'objectif de cette mise à jour est de présenter, en nous limitant aux modèles clinicopathologiques (i.e. excluant les prédicteurs génomiques), les différents types de prédicteurs ainsi que les éléments nécessaires à prendre en compte pour leur validation, étape préalable indispensable avant toute utilisation en pratique clinique. Nous présenterons ensuite les critères qui nous permettent de choisir en pratique le type de modèle et les seuils les plus adaptés. Dans une quatrième partie, nous décrirons quatre grands exemples d'utilisation de modèles prédictifs en cancérologie mammaire : I) *Adjuvant! Online* (i.e. prédiction à 10 ans de la survie globale et de la survie sans

récidive ; prédiction du bénéfice attendu des différents traitements adjuvants) ; II) la prédiction de la réponse histologique complète à la chimiothérapie néoadjuvante ; III) la prédiction du statut des ganglions non sentinelles (GNS) en cas d'envahissement du ganglion sentinelle (GS) ; et IV) la prédiction de l'envahissement d'au moins 4 ganglions axillaires en cas d'envahissement de 1 à 3 GS. Dans une cinquième partie, nous développerons les limites et précautions à considérer lors de l'utilisation de ces prédicteurs. Enfin, nous ferons le point sur les perspectives actuelles essentiellement représentées par l'utilisation combinée de prédicteurs clinicopathologiques et génomiques ainsi que le concept de prédiction spécifique.

## 1. LES DIFFÉRENTS MODÈLES PRÉDICTIFS

Un modèle prédictif ou prédicteur a pour but de modéliser l'association entre une caractéristique clinique représentée par une variable (variable à expliquer) et plusieurs variables dites variables explicatives. Il faut bien distinguer le modèle mathématique du modèle prédictif qui est l'outil qui en résulte. Les principaux modèles mathématiques sont les modèles de régression (régression logistique ou modèle de Cox) et les modèles de partition récursive. Ces modèles mathématiques permettent de créer différents types de prédicteurs. Nous nous limiterons ici aux trois plus couramment utilisés : les modèles continus (i.e. nomogrammes), les modèles binaires (scores) et les modèles de partition récursive.

### 1.1. Les modèles continus (nomogrammes)

La méthode de régression logistique est l'une des méthodes les plus importantes en statistique pour analyser la relation d'une variable par rapport à plusieurs autres. L'avantage majeur de ce modèle est qu'il permet de déterminer la probabilité de survenue de l'événement étudiée en fonction des facteurs explicatifs. De la même manière, le modèle de Cox est utilisé lorsque l'on veut modéliser le délai de survenue d'un événement (i.e. récurrence d'une maladie, apparition de métastases, décès...). La représentation graphique de ces modèles continus (i.e. donnant pour chaque patiente une probabilité de survenue de l'événement prédit par le modèle) est le nomogramme.

Le calcul de la probabilité pour chaque patiente via le nomogramme peut s'avérer long, fastidieux, peu précis et comporter un risque d'erreur. Il existe deux autres possibilités permettant de calculer la probabilité. La première est une interface Web. Sa réalisation est relativement simple eu égard à la simplicité des formules du modèle de régression logistique ou du modèle de Cox sur lequel il est basé. La seconde consiste à appliquer la formule mathématique du modèle permettant ainsi un calcul rapide et exact de la probabilité. Il faut cependant, pour appliquer cette formule, connaître la valeur des différents coefficients de régression qui ne sont pas toujours disponibles dans les différentes publications. Rappelons que l'odds ratio (OR) pour une variable correspondant à l'exponentielle du coefficient de régression. Ainsi, lorsque nous disposons des OR, il est aisé d'en déduire les coefficients de régression.

Les avantages des modèles continus sont leur simplicité, et (pour autant qu'ils soient validés sur une cohorte indépendante) leur reproductibilité. Il existe cependant plusieurs défauts à ce type de modèle. Le principal défaut est son principe même : étant donné que ce sont les pourcentages qui sont modélisés, le résultat est forcément une probabilité et non une prédiction oui/non. Les modèles linéaires ont également les défauts de leurs qualités. Plus précisément, eu égard à la simplicité du modèle, le prédicteur linéaire, et donc le modèle, ne peut être très compliqué [1]. Par exemple, on ne peut utiliser ce modèle avec des données de génomiques qui sont trop importantes en volume. Par ailleurs, le caractère linéaire du prédicteur est un autre facteur limitant car cela peut affecter sa sensibilité, c'est-à-dire sa capacité de discrimination. Enfin, le prédicteur linéaire peut être « écrasé » par un des facteurs qui est très fortement lié au pourcentage [1]. Par exemple, pour la prédiction de la survie, l'envahissement ganglionnaire écrase les autres facteurs qui, bien qu'indépendamment associés au pronostic, vont avoir en réalité peu d'influence dans le modèle car le coefficient qui leur est attribué est bien inférieur à celui de l'envahissement ganglionnaire.

## 1.2. Les scores

Les scores sont également obtenus après régression linéaire ou logistique. Les variables significatives sont incluses dans le modèle où un nombre de points leur est assigné en fonction de leur poids respectif, c'est-à-dire en fonction de la valeur du coefficient de régression de chacune d'elle. La somme des points des variables

constituant le modèle détermine un score. Un seuil est proposé par les auteurs ainsi que la probabilité de l'événement qui y est attaché. À l'inverse des nomogrammes qui fournissent une probabilité pour chaque patiente, les modèles non linéaires ont pour principal but de classer correctement les malades en fonction d'un seuil. Sur un plan pratique, un résultat oui/non peut avoir plus de valeur qu'un pourcentage à l'échelon individuel. En contrepartie, ils sont souvent moins informatifs en ne fournissant pas de quantification précise du risque [1, 2].

### 1.3. Les modèles de partition récursive

Les modèles de partition récursive peuvent être approchés comme des modèles intermédiaires aux deux premiers. Ils permettent de créer des arbres de décision qui classent les patientes ayant un niveau de risque comparable définissant ainsi des sous-groupes auxquels une probabilité est rattachée. Le choix d'un seuil peut secondairement nous ramener à un modèle binaire en regroupant tous les sous-groupes ayant une probabilité inférieure au seuil *versus* ceux ayant une probabilité supérieure au seuil [1].

## 2. VALIDATION DES MODÈLES PRÉDICTIFS

Il s'agit là d'une étape fondamentale avant l'utilisation d'un prédicteur. Tout modèle prédictif doit être validé. En effet, la qualité réelle d'un modèle ne peut pas être estimée au vu de ses résultats sur les données qui ont servi à le construire (cohorte de constitution). En fait, un modèle peut conduire à des résultats excellents (et totalement illusoire) sur les données de la cohorte de constitution, et des résultats catastrophiques sur des données indépendantes.

### 2.1. Validation interne et validation externe

Un modèle doit être validé au mieux sur une population indépendante (validation externe). En effet, elle confère au modèle ainsi testé des critères de reproductibilité et d'exportabilité. À défaut d'avoir une cohorte indépendante pour permettre une validation externe, des

méthodes de validation interne sont disponibles, basées sur des techniques de rééchantillonnage. Les trois principales techniques sont le *leave-one-out*, la validation croisée, et le *bootstrap*.

## 2.2. Paramètres de validation

La discrimination est un critère de validation classiquement décrit [3]. Cependant, elle est loin d'être suffisante pour conclure à l'exportabilité, la reproductibilité et l'utilité clinique, indispensable avant toute utilisation en pratique clinique. D'autres critères très importants sont à prendre en considération tels que la calibration, le concept d'utilité et d'efficacité clinique - qui ne peut se limiter à la seule prise en compte des taux de faux négatifs (FN) et de faux positifs (FP) - et la comparaison des résultats à un modèle optimal [2, 4, 5].

### – *Discrimination*

La discrimination est basée sur l'analyse de l'index de concordance (C index) qui est une estimation de la probabilité de concordance entre les probabilités prédites et les fréquences observées. Un paramètre équivalent est l'analyse de l'aire sous la courbe ROC (*receiver operating characteristic curve*) (AUC). Une courbe ROC est une représentation graphique du taux de FP (1-spécificité) en fonction de la sensibilité pour chaque valeur du modèle. Le calcul de l'AUC permet de mettre en évidence le pouvoir discriminant d'un test à séparer les sujets malades des non malades [3]. Lorsque l'AUC ou le C index sont supérieurs à 0,8, le pouvoir discriminant est excellent ; il est bon lorsque l'AUC est comprise entre 0,7 et 0,8 [6]. Une AUC à 0,5 correspond à une prédiction au hasard. Il est toujours important de connaître l'intervalle de confiance (à 95 % le plus souvent) de ces paramètres estimant la discrimination. Pour les données censurées (modèle de Cox), seul le C index est utilisable [3, 7].

### – *Calibration*

La calibration est un paramètre central à prendre en compte. En effet, un modèle discriminant mais mal calibré aura une utilité clinique très limitée.

La calibration ayant pour but de comparer les probabilités prédites par le modèle et les fréquences réellement observées, elle n'a bien sûr d'intérêt que pour les modèles continus.

Elle est étudiée par une représentation graphique, la courbe de calibration qui a pour fonction :  $y = \beta x + \alpha$  où  $\alpha$  et  $\beta$  (en fait les  $\beta_i$  de

chaque variable du modèle) sont estimés par le modèle de régression logistique (axe des x = probabilités prédites et axe des y = fréquences observées). L'ordonnée à l'origine (intercept)  $\alpha$  est l'intersection entre la droite ainsi définie et l'axe des abscisses, c'est-à-dire la valeur de y quand x = 0.

Un modèle est correctement calibré lorsque les probabilités prédites sont statistiquement non significatives par rapport aux fréquences observées. La droite idéale est alors  $y = x$ , c'est-à-dire  $\beta = 1$  et  $\alpha = 0$ . Tester la calibration d'un modèle revient à comparer la droite idéale,  $y = x$  et la courbe de calibration  $y = \beta x + \alpha$ , c'est-à-dire tester le likelihood ratio de la double hypothèse nulle :  $\alpha = 0$  et  $\beta = 1$ . Ce test est un  $\chi^2$  à 2 degrés de liberté. Un modèle est correctement calibré lorsque les deux droites sont non statistiquement différentes ( $p > 0,05$ ) [2, 4, 5, 7, 8].

Deux paramètres importants de la calibration sont à analyser. Il s'agit des erreurs moyenne (E moy) et maximale (E max). Ces 2 paramètres ont un intérêt clinique important et ce d'autant qu'un modèle continu n'est pas correctement calibré. En effet, E moy et E max donnent une idée de l'intervalle de confiance lors de la prédiction individuelle [2, 5]. Prenons comme exemple un nomogramme prédictif de l'envahissement des GNS en cas de GS métastatique. Ce nomogramme n'est pas parfaitement calibré et il a une E max à 6 %. Lorsque le modèle prédit pour une patiente une probabilité d'envahissement des GNS de 10 %, alors cette probabilité va varier entre 4 % et 16 %. Ainsi, la connaissance et la prise en compte de E moy et E max nous paraissent indispensables pour l'utilisation en pratique des modèles continus (nomogramme).

### – *Concept d'utilité et d'efficacité clinique*

Le but d'un prédicteur est d'être utile et efficace en pratique courante. En effet, un modèle discriminant parfaitement calibré et validé peut avoir une utilité clinique limitée pour la pratique. Cette notion d'utilité clinique d'un modèle est difficile à appréhender et à modéliser et va dépendre en grande partie de la question à laquelle il doit répondre [2, 5]. Schématiquement, il repose sur :

**I).** Les qualités intrinsèques du modèle (sensibilité, spécificité), les valeurs prédictives positive (VPP) et négative (VPN) et les taux de FN et de FP. Par exemple, pour les modèles prédictifs du non-envahissement des GNS en cas de GS positif, le paramètre important est un taux de FN le plus faible possible (un taux de FN inférieur ou égal à 5 % est considéré comme acceptable par les auteurs en



correspondant au taux de FN du curage axillaire). Dans d'autre cas, le modèle devra privilégier un taux de FP le plus bas possible.

II). La notion d'efficacité du modèle repose sur sa capacité à sélectionner un nombre suffisant de patientes. En prenant toujours comme exemple les modèles prédictifs de non-envahissement des GNS en cas de GS métastatique, 40 à 70 % des patientes n'auront pas d'envahissement des GNS. Ainsi, la proportion de patientes sélectionnées par le modèle comme appartenant au groupe à bas risque est un paramètre clé. En effet, un modèle discriminant, correctement calibré, avec un taux de FN acceptable mais qui ne sélectionnerait que quelques pourcents de la population verrait son intérêt clinique très limité [2, 5].

– **Comparaison à un modèle optimal**

Coutant *et al.* ont proposé de comparer les résultats obtenus en termes de discrimination (AUC ou C index) pour les modèles testés à celui du modèle de régression logistique (ou de modèle de Cox) réalisé sur la population de validation [2, 5]. Ainsi, ce modèle peut être considéré comme un véritable modèle optimal, c'est-à-dire offrant par là-même les meilleurs résultats possibles sur cette population. Il consiste à réaliser une régression logistique (ou un modèle de Cox) sur la population de validation. La réalisation d'un *bootstrapping* permet d'obtenir une estimation la moins biaisée possible de l'AUC. Le résultat de l'AUC (ou du C index) ainsi obtenu doit être considéré comme le maximum qui peut être observé sur cette population [2, 5].

### 3. UTILISATION EN PRATIQUE D'UN PRÉDICTEUR : QUELS TYPES DE MODÈLES UTILISER ? QUELS SEUILS CHOISIR ?

#### 3.1. Quels types de modèles utiliser ?

L'une des questions clés lorsque l'on souhaite utiliser ou développer un prédicteur est le choix du type de modèle le plus adapté (i.e. modèle continu, modèle binaire, modèle de partition récursive). Ce choix va résulter d'un compromis entre, d'une part, les qualités et défauts propres à chaque type de modèles et, d'autre part, la question posée, c'est-à-dire la variable à expliquer que le modèle est censé

modéliser. Ainsi, dans certain cas souhaitera-t-on privilégier un taux le plus faible possible de FN (par exemple pour les modèles prédictifs du non-envahissement des GNS en cas de GS positifs), alors que dans d'autres cas, le modèle devra minimiser au maximum le risque de FP (par exemple pour un prédicteur qui conduirait à proposer une intervention chirurgicale ou un traitement associé à un taux de morbidité non négligeable).

Plusieurs études ont comparé ces différents types de modèles. Kattan *et al.* et Shariat *et al.* ont montré sur plusieurs séries de patients ayant un cancer urologique (prostate essentiellement) que les modèles continus étaient comparables ou supérieurs aux modèles de partition récursive, aux réseaux de neurones artificiels et aux autres méthodes de « *machine learning* » [9-12]. Rouzier *et al.* ont comparé deux types de prédicteurs de la réponse complète à la chimiothérapie néoadjuvante à base d'anthracyclines dans les cancers du sein : un modèle de régression logistique (nomogramme) et un modèle de partition récursive. Ces deux prédicteurs ont été créés sur la même cohorte de constitution et validés sur la même cohorte indépendante. Les résultats étaient fortement en faveur du modèle continu, en accord avec ceux rapportés par Kattan *et al.* et Shariat *et al.* dans les cancers urologiques [1].

Cependant, il n'est pas possible de généraliser ce type de résultats car le choix du modèle le plus adapté va dépendre en grande partie de la question posée (variable à expliquer), des caractéristiques intrinsèques des variables explicatives et de leur relation. Par exemple, les modèles continus sont théoriquement supérieurs au modèle de régression logistique lorsque la véritable relation entre les variables explicatives et la variable à expliquer est linéaire. Dans le cas inverse, les modèles de partition récursive peuvent s'avérer nettement plus performants [1, 2, 5].

### 3.2. Quels seuils choisir ?

Un modèle continu, en donnant pour une patiente une probabilité de présence de l'événement prédit, peut être approché de deux façons :

**I).** Il fournit pour chaque patiente une estimation quantitative du risque. Cette estimation sera d'autant plus précise que le modèle sera correctement calibré. Elle sera pondérée par les paramètres de calibration (E moy et E max) [2].

II). Cependant, pour prendre une décision, il est souvent utile de disposer d'un seuil, c'est-à-dire de passer d'un modèle continu à un modèle binaire. En prenant comme exemple le nomogramme du MSKCC prédictif de l'invasion des GNS en cas de GS métastatique, le choix du seuil est crucial pour décider si un curage axillaire peut être évité ou non. Ce seuil est le plus souvent choisi de manière arbitraire et empirique. Dans l'exemple présent, les auteurs ont proposé deux seuils (5 % et 10 %) qu'ils ont arbitrairement choisis. Par la suite, dans les différentes études de validations, le seuil de 10 % a été retenu.

Une autre approche consiste à identifier puis comparer différents seuils permettant d'optimiser la prédiction, appelés « *Optimal Threshold* ». Les différents logiciels de statistique permettent de définir ces seuils, tel que le package « *Presence-Absence* » du logiciel R (<http://lib.stat.cmu.edu/R/CRAN/>).

Différents seuils optimaux existent, privilégiant soit la sensibilité, la spécificité, la proportion de sujets correctement classés (PCC), ou encore le coefficient de Kappa... Le choix du type de seuil optimal est intimement lié à la question posée que le prédicteur est censé modéliser. Le but ici n'est pas de décrire ces différents seuils. Citons simplement les quatre plus importants : le point MaxPCC, le point de Youden, le point de MaxKappa, et le point MinROCdist [13-16].

Finalement, les modèles continus semblent représenter le prédicteur le plus polyvalent. Non seulement ils fournissent une quantification individuelle du risque avec un intervalle de confiance que l'on peut approcher par E max et E moy de la courbe de calibration mais ils permettent aussi une prédiction oui/non très utile pour l'aide à la décision. Le passage du modèle continu au modèle binaire sera d'autant plus fiable et pertinent qu'il existe plusieurs méthodes pour définir le seuil optimal, en fonction des caractéristiques des cohortes (prévalence par exemple) mais aussi de la question posée et des qualités intrinsèques du modèle que l'on souhaite favoriser (taux de FN faible, taux de FP faible...).

## 4. LES DIFFÉRENTS MODÈLES PRÉDICTIFS CLINICOPATHOLOGIQUES EN CANCÉROLOGIE MAMMAIRE (TABLEAU 1)

Tableau 1 - Prédicteurs clinicopathologiques en cancérologie mammaire

Prédicteurs	Auteurs, année	Type de modèle	Existence d'une validation indépendante	Modèle validé	Utilisation en routine	Interface Web
<b>Adjuvant! Online</b>						
Version 8.0	Ravdin <i>et al.</i> , 2001 [17]	Continu / Rlog	Olivotto <i>et al.</i> , 2005 [18]	oui	oui	Oui*
<b>Réponse à la chimiothérapie néoadjuvante</b>						
Prédicteur de réponse histologique complète Néoadjuvant!	Rouzier <i>et al.</i> , 2005 [8]	Continu / Rlog Continu / Cox	Rouzier <i>et al.</i> , 2005 [8]	oui oui	oui oui	Oui**
<b>Statut des GNS en cas d'envahissement du GS</b>						
MSKCC nomogramme	Van Zee <i>et al.</i> , 2003 [42]	Continu / Rlog	Coutant <i>et al.</i> , 2009 [2]	oui	non	Oui#
Mayo nomogramme	Degnim <i>et al.</i> , 2005 [41]	Continu / Rlog		non	non	
Cambridge nomogramme	Pal <i>et al.</i> , 2008 [45]	Continu / Rlog		non	non	
Stanford nomogramme	Kohrt <i>et al.</i> , 2008 [44]	Continu / Rlog		non	non	Oui#
Score de Tenon	Barranger <i>et al.</i> , 2005 [40]	score / Rlog		oui	non	
Score du MDA	Hwang <i>et al.</i> , 2003 [36]	score / Rlog		oui	non	
Score de Saidi	Saidi <i>et al.</i> , 2004 [46]	score / Rlog		non	non	
Modèle de Kohrt	Kohrt <i>et al.</i> , 2008 [44]	Partition récursive		non	non	
<b>Envahissement d'au moins 4 ganglions axillaires en cas d'envahissement de 1 à 3 GS</b>						
Katz nomogramme	Katz <i>et al.</i> , 2008[51]	Continu / Rlog	Werkoff <i>et al.</i> , 2009 [5]	oui	non	
Score de Louisville excluant la MD	Chapgar <i>et al.</i> , 2007[50]	score / Rlog		oui	non	
Score de Louisville incluant la MD		score / Rlog		oui	non	
GNS : ganglion non sentinelle ; GS : ganglion sentinelle ; MSKCC : Memorial Sloan-Kettering Cancer Center ; MDA : M.D. Anderson Cancer Center ; MD : méthode de détection histologique des GS ; Rlog : modèle de régression logistique ; Cox : modèle de Cox * <a href="http://www.adjuvantonline.com">www.adjuvantonline.com</a> ** <a href="http://www3.mdanderson.org/app/medcalc/index.cfm?pagename=jsconvert2">http://www3.mdanderson.org/app/medcalc/index.cfm?pagename=jsconvert2</a> # <a href="http://www.mskcc.org/mskcc/html/5794.cfm">http://www.mskcc.org/mskcc/html/5794.cfm</a> ## <a href="http://www-stat.stanford.edu/~olshen/NSLcalculator">http://www-stat.stanford.edu/~olshen/NSLcalculator</a>						

## 4.1. Adjuvant! Online

Ravdin *et al.* ont développé en 2001 un prédicteur - *Adjuvant! Online* ([www.adjuvantonline.com](http://www.adjuvantonline.com)) - permettant de déterminer à 10 ans la survie globale et la survie sans récurrence pour une patiente à partir de 6 variables clinico-pathologiques (i.e. l'âge de la patiente, la taille de la tumeur, le statut ganglionnaire (nombre de ganglions positifs), le grade histologique, le statut vis-à-vis des récepteurs aux œstrogènes (statut ER) et les éventuels facteurs de comorbidités). Il permet également de prédire le bénéfice attendu des différents traitements adjuvants [17]. Ce modèle a été validé par Olivotto *et al.* sur une cohorte canadienne indépendante de 4 083 patientes [18]. Largement utilisé, on estime que 75 % des oncologues américains le consultent [19, 20] et il a été montré qu'il influençait largement le choix des patientes en pratique clinique [20, 21]. Une première remarque est que les prédictions du programme *Adjuvant! Online* sont basées sur les données de mortalité de cancer du sein rapportées dans la SEER (*Surveillance Epidemiology and End Results*) database et sur les résultats des essais randomisés comparant chimiothérapie *versus* pas de chimiothérapie adjuvante. Ainsi, le bénéfice de la chimiothérapie adjuvante prise en compte par le programme n'est pas réellement établi en fonction de l'ensemble des caractéristiques cliniques et biologiques.

Ozanne *et al.* ont étudié la sensibilité du prédicteur aux entrées en termes de comorbidités [20]. Les auteurs ont montré que les prédictions de la mortalité à 10 ans dépendaient largement des variables « nombre de ganglions positifs » et « comorbidités » (i.e. aucune, comorbidités mineures, en accord avec l'âge, et trois critères de gravité croissante pour comorbidités majeures), en particulier chez les femmes âgées. La variable « comorbidités » était le critère le plus influent sauf pour les femmes de moins de 40 ans. Or, parmi les différents critères demandés, c'est sans doute le plus difficile à renseigner, et ce d'autant qu'il n'existe pas d'évaluation standardisée pour ce paramètre. Ces résultats montrent combien une mauvaise évaluation des comorbidités risque de conduire à une sur- ou sous-estimation du risque (et donc à un risque de sur- ou sous-traitement) [20]. Nous verrons dans le chapitre 5 d'autres limites et précautions à considérer pour l'utilisation de ce prédicteur.

## 4.2. Modèles prédictifs de la réponse à la chimiothérapie néoadjuvante

Le concept de la chimiothérapie néoadjuvante est séduisant car il permet de connaître la chimiosensibilité d'une tumeur « *in vivo* » [22-24] et de proposer ainsi un traitement « à la carte ». En cas de réponse histologique complète, les traitements adjuvants peuvent être évités. Cette hypothèse est supportée par plusieurs études confirmant le très bon pronostic à long terme des patientes ayant une réponse histologique complète [22, 25-28]. À l'inverse, en cas de réponse insuffisante, d'autres drogues sans résistance croisée peuvent être utilisées [29]. La réponse d'une tumeur à un agent cytotoxique est bien sûr fortement liée aux caractéristiques intrinsèques propres à chaque tumeur et il est possible d'établir une prédictivité de l'efficacité de cette chimiothérapie.

Différentes variables clinicopathologiques ont été identifiées comme prédictives de la réponse à la chimiothérapie néoadjuvante (statut ER-négatif, HER-2 surexprimé, grade histologique élevé) [8, 30, 31]. Rouzier *et al.* ont développé un nomogramme basé sur un modèle de régression logistique de réponse à la chimiothérapie néoadjuvante à base d'anthracyclines incluant comme variables l'âge de la patiente, la taille tumorale (T), le grade histologique, le statut ER, et le nombre de cures de chimiothérapie [8]. Basé sur un modèle de Cox, les auteurs ont également développé un prédicteur permettant de prédire la survie sans récidence à 5 ans et à 10 ans chez les patientes ayant reçu une chimiothérapie néoadjuvante [8]. Une interface Web appelée *Neo!adjuvant* a été développée par les auteurs et est disponible sur les Web sites du M.D. Anderson Cancer Center : <http://www.mdanderson.org> (<http://www3.mdanderson.org/app/medcalc/index.cfm?pagename=jsconvert2>).

## 4.3. Modèles prédictifs du statut des GNS en cas d'envahissement du GS

Être capable de prédire l'absence de maladie résiduelle dans les GNS en cas d'envahissement des GS permettrait d'éviter un curage axillaire chez 40 à 70 % des patientes [2, 4, 32-39]. Des facteurs prédictifs d'envahissement des GNS en cas de GS métastatiques ont été mis en évidence dans de nombreuses études [33, 34, 36, 39-43]. Cependant, les performances d'un facteur prédictif pris isolément sont insuffisantes, y compris pour ceux apparemment associés à un très faible risque d'envahissement des GNS (micrométastase et a fortiori

cellules isolées, taille tumorale inférieure à 10 mm, absence de rupture capsulaire, GS métastatique unique) [2, 34, 36, 40-42]. Dans ce contexte, l'utilisation de prédicteurs clinicopathologiques peut s'avérer très utile pour sélectionner les patientes ne nécessitant pas de curage axillaire complémentaire.

Neuf modèles ont été à ce jour développés afin de prédire le risque d'envahissement des GNS en cas de GS métastatique : quatre nomogrammes développés par Van Zee *et al.* (Memorial Sloan-Kettering Cancer Center (MSKCC) nomogramme) ; Degnim *et al.* (Mayo nomogramme) ; Pal *et al.* (Cambridge nomogramme) ; et Kohrt *et al.* (Stanford nomogramme) ; trois scores (le score de Tenon, le score du M.D. Anderson Cancer Center (MDA score), et le score proposé par Saidi *et al.*) ; et deux modèles de partitions récursives développés par Kohrt *et al.* [36, 40-42, 44-46].

Les méthodologies des premières validations rapportées pour ces modèles étaient très hétérogènes. En effet, seulement 4 de ces 9 modèles (le nomogramme du MSKCC, le nomogramme de Stanford, le score du MDA et le score de Tenon) ont été évalués sur une population indépendante. Le nomogramme du MSKCC a été évalué par 17 équipes [2, 4, 41, 42, 44, 47, 48] avec des AUC qui variaient de 0,58 à 0,86. Il a été validé par douze équipes (AUC > 0,7) [2, 4, 41, 42, 44, 47, 48], et non validé par 3 autres équipes dont Alran *et al.* pour le sous-groupe de patientes dont le GS ne contenait que des micrométastases ou des cellules tumorales isolées [2, 4, 47]. Quatre équipes seulement ont étudié la calibration de ce modèle sur leur population [2, 4]. Seulement trois équipes ont rapporté leur taux de FN [2, 4, 41, 48]. Le score du MDA a été validé par Dauphine *et al.* avec une AUC à 0,7 mais sans information sur le taux de FN [49]. À l'inverse, Ponzone *et al.* rapportaient un taux de FN de 18,8 % ne validant pas le score mais sans fournir l'AUC [48]. Kohrt *et al.* ont validé sur une population indépendante de 77 patientes leur modèle (nomogramme de Stanford) et le nomogramme du MSKCC avec des AUC de 0,74 et 0,62 respectivement sans rapporter ni la calibration ni le taux de FN [44].

Ces modèles ont récemment été validés sur une cohorte prospective, multicentrique et indépendante de 561 patientes ayant au moins un GS métastatique avec une méthodologie robuste [2]. En prenant en compte les différents critères de validation (i.e. discrimination, concept d'utilité clinique basé sur le taux de FN et la proportion de patientes sélectionnées par le modèle, la comparaison avec le modèle optimal dont l'AUC était de 0,81 et - pour les modèles continus - la calibration), les deux prédicteurs les plus performants

étaient le nomogramme du MSKCC et le score de Tenon. Des résultats similaires ont été rapportés pour le sous-groupe de patientes dont le GS ne contenait que des micrométastases ou des cellules isolées [2].

Dans cette étude, les auteurs montrent combien la calibration est importante pour juger de la pertinence d'un modèle. En effet, un des nomogrammes évalués avait une AUC tout à fait correcte à 0,72 (IC 95 % : 0,7-0,74) mais était non calibré avec une E Moy à 23 %, le rendant totalement inutilisable sur leur cohorte. Un des autres nomogrammes avait une AUC comparable (0,72 ; IC 95 % : 0,7-0,75) mais était parfaitement calibré (E Moy à 0,03 %). Ne prendre en compte que les valeurs de l'AUC conduisait à une conclusion totalement erronée quant à l'équivalence entre ces 2 modèles [2].

#### 4.4. Modèles prédictifs de l'envahissement d'au moins 4 ganglions axillaires en cas d'envahissement de 1 à 3 GS

Trois modèles prédictifs de l'envahissement d'au moins 4 ganglions axillaires chez les femmes ayant un cancer du sein avec envahissement de 1 à 3 GS ont été rapportés. Deux modèles sont des scores développés par Chapgar *et al.* (Score de Louisville scores excluant ou incluant la méthode de détection) et le dernier modèle est un nomogramme développé par Katz *et al.* [50, 51]. La prédiction de l'envahissement d'au moins 4 ganglions axillaires a un intérêt au regard des indications de radiothérapie de paroi adjuvante. Les recommandations actuelles de « l'*American Society of Clinical Oncology* » (ASCO) [52] et de « l'*American Society of Therapeutic Radiology and Oncology* » [53] préconisent la réalisation d'une radiothérapie de paroi postmastectomie en cas de tumeur du sein mesurant plus de 5 cm et/ou lorsqu'au moins 4 ganglions axillaires sont envahis [52, 53]. Sélectionner les patientes ne justifiant pas d'une radiothérapie postopératoire pourrait leur permettre de bénéficier d'une reconstruction mammaire immédiate [5, 50, 51]. Werkoff *et al.* ont réalisé une validation de ces 3 modèles sur une cohorte prospective, multicentrique et indépendante de 536 patientes ayant entre 1 et 3 GS métastatiques. Les trois modèles avaient des performances satisfaisantes en termes de discrimination et de taux de FN. Cependant, le nomogramme de Katz *et al.* était supérieur sur le plan de l'efficacité clinique. De plus, il était parfaitement calibré (E max à 1 %) [5].



## 5. LIMITES ET PRÉCAUTIONS LORS DE L'UTILISATION DES PRÉDICTEURS

Une des limites pour l'utilisation d'un prédicteur peut être sa mauvaise exportabilité. Par exemple, le prédicteur *Adjuvant! Online* a été élaboré sur des cohortes nord-américaines (États-Unis pour la cohorte de constitution et canadienne pour la cohorte de validation) [17, 18]. Or les caractéristiques démographiques et clinicopathologiques peuvent parfois être très différentes entre les différents pays. Ainsi, un prédicteur fiable sur des cohortes ayant des caractéristiques similaires à celles des cohortes de constitution et de validation peut aboutir à des prédictions erronées sur d'autres cohortes avec des caractéristiques différentes. Ceci a été parfaitement rapporté par Campbell *et al.* [54]. Les auteurs ont analysé les performances du prédicteur *Adjuvant! Online* sur une cohorte de 1 065 patientes prises en charge en Grande-Bretagne entre 1986 et 1996. Les auteurs ont montré que le prédicteur surestimait constamment les taux de survie global (de 5,54 %,  $p < 0,001$ ), et de survie sans récurrence (de 3,51 %,  $p = 0,001$ ). Ces différences persistaient dans les différentes analyses de sous-groupes (en fonction des données démographiques, de l'âge, des caractéristiques clinicopathologiques ou des traitements adjuvants reçus). Une des explications rapportées par les auteurs est que les taux de mortalité par cancer du sein sont constamment plus bas aux États-Unis comparés à ceux de la Grande-Bretagne [54]. Ainsi ces résultats semblent montrer qu'il est nécessaire de valider ces prédicteurs dans les différents pays avant leur utilisation en pratique, voire de les optimiser en fonction des caractéristiques propres à chaque pays.

Une autre limite tient du fait que les caractéristiques clinicopathologiques des cohortes évoluent au cours du temps, de même que les protocoles de traitements. Par exemple, les prédicteurs de réponse à la chimiothérapie néoadjuvante développés par Rouzier *et al.* ont été élaborés à partir de cohortes de patientes ayant reçu au maximum 4 cycles d'anthracyclines ou encore ayant une surexpression de HER-2 sans traitement par trastuzumab [8]. Or actuellement, des protocoles incluant 6 cycles d'anthracyclines sont développés. De même, les patientes ayant une surexpression de HER-2 reçoivent du trastuzumab avec des résultats supérieurs en termes de réponse et de survie. Appliqué à ces nouvelles cohortes de patientes, le prédicteur va nécessairement sous-estimer le taux de réponse à la chimiothérapie néoadjuvante et les taux de survie. La même remarque est vraie pour le prédicteur *Adjuvant! Online*. Bien que plusieurs versions implémentées

aient été développées (actuellement version 8.0), le prédicteur ne prend toujours pas en compte les thérapeutiques ciblées telles que le trastuzumab pour les patientes surexprimant HER-2 [17].

Ces exemples montrent combien une implémentation et une optimisation régulière de ces différents prédicteurs sont indispensables afin qu'ils conservent leur pertinence.

## 6. PERSPECTIVES

Parallèlement, l'essor de la génomique à haut débit a permis une caractérisation moléculaire plus globale, détaillée et objective de la maladie - la classification moléculaire des cancers du sein en est un exemple - et a également permis le développement de véritables signatures génomiques pronostiques ou prédictives de la réponse à la chimiothérapie néoadjuvante [22, 28, 55-68].

L'une des perspectives en termes de prédiction concerne indiscutablement la prédiction combinée. Plusieurs auteurs ont montré que l'utilisation combinée de prédicteurs clinicopathologiques et génomiques permettait d'optimiser leurs valeurs prédictives [28, 69]. Stephenson *et al.* ont étudié les différences de profils d'expression géniques entre 37 tissus tumoraux issus de patients ayant un cancer de la prostate avec récurrence et 42 échantillons de patients sans récurrence [69]. Deux modèles de régression logistique ont été construits pour prédire la récurrence après prostatectomie radicale : un modèle basé uniquement sur la différence de profil d'expression génique entre les 2 classes et un deuxième modèle intégrant à la fois le profil d'expression génique et un nomogramme validé et basé sur des variables clinicopathologiques et biologiques. La pertinence de ces modèles était testée par validation interne selon la technique *leave-one-out*. Le modèle basé sur l'expression génique avait un C index de 0,75. Ce résultat était inférieur à la pertinence du nomogramme qui avait un C index de 0,84 ( $p = 0,001$ ). Le modèle combiné avait un C index de 0,89, ce qui était significativement meilleur que le nomogramme ( $p = 0,009$ ) et que le modèle basé sur le profil d'expression génique seul ( $p = 0,001$ ). Le modèle combiné améliorerait surtout la prédiction pour les patients qui avaient une probabilité de récurrence indéterminée par le nomogramme. Les auteurs concluaient que l'approche combinée améliorerait significativement la prédiction de récurrence de cancer de la prostate par rapport aux modèles cliniques et génomiques seuls [69].

Des résultats similaires ont été rapportés pour la prédiction de la réponse histologique complète à la chimiothérapie néoadjuvante dans le cancer du sein. Rouzier *et al.* ont combiné deux prédicteurs : leur prédicteur clinicopathologique (nomogramme) et un prédicteur génomique (*DLDA30-gene*). Les performances de prédiction avec le prédicteur clinicopathologique ou le prédicteur génomique étaient de 84 %. En cas d'utilisation combinée de ces deux prédicteurs, la performance augmentait à 91 % et 97 % selon le modèle utilisé (régression logistique et partition récursive, respectivement).

Le concept de prédiction combinée peut aussi intéresser deux prédicteurs clinicopathologiques et être appliqué à la problématique de la prédiction du statut des GNS en cas d'invasion du GS. En effet, à côté de l'augmentation des performances prédictives que permet la prédiction combinée, l'utilisation combinée d'un modèle continu et d'un modèle binaire peut s'avérer très utile pour la pratique clinique : un résultat binaire (positif/négatif) est très pertinent sur le plan pratique alors qu'une quantification du risque est plus informative.

Le développement de modèles permettant de prédire spécifiquement le risque de métastases en fonction du site (par exemple un prédicteur spécifique du risque de métastases cérébrales, ou du risque de métastases osseuses) constitue également une perspective intéressante. En effet, ces modèles pourraient être utiles pour proposer des traitements spécifiques tels que les biphosphonates pour les patientes à risque de métastases osseuses ou encore un plan personnalisé de surveillance radiologique.

Nous sommes en train actuellement de développer un modèle permettant de prédire spécifiquement le risque de métastases cérébrales chez les patientes ayant un cancer du sein métastatique autre que cérébrale. Le rationnel d'un tel modèle (en identifiant un sous-groupe de patientes à risque de développer des métastases cérébrales) est qu'il permettrait de sélectionner les patientes pour une irradiation prophylactique de l'encéphale en totalité (IPET). Actuellement, une IPET est recommandée chez les patients ayant un cancer du poumon à petites cellules [70, 71] et non à petites cellules [72]. Ce traitement prophylactique a fait la preuve de son efficacité chez ces patients avec une augmentation significative de la survie, une diminution du taux de récurrence cérébrale et une amélioration de la qualité de vie [70-72]. Cependant, sa morbidité n'est pas négligeable. Noel *et al.* ont rapporté une incidence de complication après irradiation de l'encéphale en totalité de 11 % [73] ; la complication la plus fréquente étant l'encéphalopathie radique suivie parfois par une démence [73]. Ainsi,

une IPET chez des patientes non sélectionnées, c'est-à-dire non identifiées comme appartenant à un haut niveau de risque de métastases cérébrales, ne semble pas acceptable tant sur un plan éthique (traitement inutile dans 85 à 95 % des cas avec un risque de complications graves estimé à 11 %) que médico-économique et organisationnel.

Enfin, ces prédicteurs spécifiques pourraient être utilisés afin de rationaliser le schéma d'un essai contrôlé. Reprenons l'exemple du nomogramme prédictif de la survenue de métastases cérébrales en cours de développement et d'un essai contrôlé d'IPET. Ne sélectionner dans le cadre de cet essai contrôlé que les patientes à haut risque de développer une métastase cérébrale permettrait d'augmenter l'efficacité de l'essai (diminution du nombre d'inclusions tout en conservant la même puissance, diminution des coûts et des durées) tout en diminuant la morbidité (minoration du nombre d'irradiation de l'encéphale en totalité chez les patientes qui ne feraient pas de métastases cérébrales).

## CONCLUSION

L'utilisation de prédicteurs clinicopathologiques offre la possibilité de proposer aux patientes une information pertinente et objective, mais surtout personnalisée tant sur le plan pronostique (risque de récurrence) que prédictif (réponses aux traitements adjuvants ou à la chimiothérapie néoadjuvante). Cependant, une validation rigoureuse de ces prédicteurs est indispensable avant toute utilisation en pratique clinique. Les paramètres à prendre en compte sont la discrimination, la calibration, le concept d'utilité et d'efficacité clinique et la comparaison à un modèle optimal.

Insistons également sur le fait que les caractéristiques démographiques et clinicopathologiques (découverte de nouveaux biomarqueurs par exemple) ainsi que les protocoles de traitements - néoadjuvants ou adjuvants - évoluent rapidement. Ainsi, ces différents prédicteurs doivent être régulièrement implémentés et optimisés afin de conserver leur pertinence.

Les perspectives sont essentiellement représentées par la prédiction spécifique et la prédiction combinée qui apparaissent de plus en plus comme des éléments incontournables dans le domaine de la prédiction en cancérologie mammaire.

## Bibliographie

- [1] Rouzier R, Coutant C, Lesieur B et al. Direct comparison of logistic regression and recursive partitioning to predict chemotherapy response of breast cancer based on clinical pathological variables. *Breast Cancer Res Treat* 2009 Sep;117(2):325-31. Epub 2009 Jan 17.
- [2] Coutant C, Olivier C, Lambaudie E et al. Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study. *J Clin Oncol* 2009;27:2800-2808.
- [3] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36.
- [4] Coutant C, Rouzier R, Olivier C et al. [Models to predict non sentinel lymph node status in breast cancer patients with metastatic sentinel lymph node]. *Gynecol Obstet Fertil* 2009;37:160-166.
- [5] Werkoff G, Lambaudie E, Fondrinier E et al. Prospective multicenter comparison of models to predict four or more involved axillary lymph nodes in patients with breast cancer with ont to three metastatic sentinel lymph nodes. *J Clin Oncol* 2009; in press.
- [6] Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-1293.
- [7] Cox D. Two further applications of a model for binary regression. *Biometrika* 1958;45:562-565.
- [8] Rouzier R, Pusztai L, Delaloge S et al. Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. *J Clin Oncol* 2005;23:8331-8339.
- [9] Kattan MW. Comparison of Cox regression with other methods for determining prediction models and nomograms. *J Urol* 2003;170:S6-9;discussion S10.
- [10] Kattan MW. Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. *Curr Opin Urol* 2003;13:111-116.
- [11] Shariat SF, Capitanio U, Jeldres C, Karakiewicz PI. Can nomograms be superior to other prediction tools? *BJU Int* 2009;103:492-495;discussion 495-497.
- [12] Shariat SF, Karakiewicz PI, Suardi N, Kattan MW. Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. *Clin Cancer Res* 2008;14:4400-4407.
- [13] Allouche O. Assessing the accuracy of species distribution models: prevalence, kappa, and the true skill statistic (TSS). *Journal of Applied Ecology* 2006;43:1223-1232.
- [14] Cantor SB, Sun CC, Tortolero-Luna G et al. A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. *J Clin Epidemiol* 1999;52:885-892.
- [15] Freeman E, Moisen D. PresenceAbsence: an R package for presence-absence analysis. *Journal of statistical Software* 23 2008;11.
- [16] Manel S. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 2001; 38:921-931.
- [17] Ravdin PM, Siminoff LA, Davis GJ et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001;19:980-991.
- [18] Olivetto IA, Bajdik CD, Ravdin PM et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 2005;23:2716-2725.
- [19] Love N. Management of breast cancer in the adjuvant and metastatic settings. *Patterns of Care* 2 2005.
- [20] Ozanne EM, Braithwaite D, Sepucha K et al. Sensitivity to input variability of the Adjuvant! Online breast cancer prognostic model. *J Clin Oncol* 2009;27:214-219.
- [21] Peele PB, Siminoff LA, Xu Y, Ravdin PM. Decreased use of adjuvant breast cancer therapy in a randomized controlled trial of a decision aid with individualized risk information. *Med Decis Making* 2005;25:301-307.
- [22] Bianchini G, Gianni L. Introducing new molecular technologies into routine clinical cancer care: is there an impact on the treatment of breast cancer? *Ann Oncol* 2008;19 Suppl 7:vii177-183.
- [23] Kaufmann M, Hortobagyi GN, Goldhirsch A et al. Recommendations from an international expert panel on the use of neoadjuvant (primary)

systemic treatment of operable breast cancer: an update. *J Clin Oncol* 2006;24:1940-1949.

[24] Mauri D, Pavlidis N, Ioannidis JP. Neoadjuvant versus adjuvant systemic treatment in breast cancer: a meta-analysis. *J Natl Cancer Inst* 2005;97:188-194.

[25] Bonadonna G, Valagussa P, Brambilla C et al. Primary chemotherapy in operable breast cancer: eight-year experience at the Milan Cancer Institute. *J Clin Oncol* 1998;16:93-100.

[26] Fisher B, Bryant J, Wolmark N et al. Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. *J Clin Oncol* 1998;16:2672-2685.

[27] Gralow JR, Burstein HJ, Wood W et al. Preoperative therapy in invasive breast cancer: pathologic assessment and systemic therapy issues in operable disease. *J Clin Oncol* 2008;26:814-819.

[28] Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 2009;360:790-800.

[29] Dieras V, Fumoleau P, Romieu G et al. Randomized parallel study of doxorubicin plus paclitaxel and doxorubicin plus cyclophosphamide as neoadjuvant treatment of patients with breast cancer. *J Clin Oncol* 2004;22:4958-4965.

[30] Chang J, Powles TJ, Allred DC et al. Biologic markers as predictors of clinical outcome from systemic therapy for primary operable breast cancer. *J Clin Oncol* 1999;17:3058-3063.

[31] Ellis P, Smith I, Ashley S et al. Clinical prognostic and predictive factors for primary chemotherapy in operable breast cancer. *J Clin Oncol* 1998;16:107-114.

[32] Chu KU, Turner RR, Hansen NM et al. Do all patients with sentinel node metastasis from breast carcinoma need complete axillary node dissection? *Ann Surg* 1999;229:536-541.

[33] Chu KU, Turner RR, Hansen NM et al. Sentinel node metastasis in patients with breast carcinoma accurately predicts immunohistochemically detectable nonsentinel node metastasis. *Ann Surg Oncol* 1999;6:756-761.

[34] Coutant C, Morel O, Antoine M et al. [Is axillary lymph node dissection always necessary in breast cancer patients with a positive sentinel node?]. *J Chir (Paris)* 2007;144:492-501.

[35] Coutant C, Rouzier R, Fondrinier E et al. Validation of the Tenon breast cancer score for predicting non-sentinel lymph node status in

breast cancer patients with sentinel lymph node metastasis: a prospective multicenter study. *Breast Cancer Res Treat* 2009;113:537-543.

[36] Hwang RF, Krishnamurthy S, Hunt KK et al. Clinicopathologic factors predicting involvement of nonsentinel axillary nodes in women with breast cancer. *Ann Surg Oncol* 2003;10:248-254.

[37] Kamath VJ, Giuliano R, Dauway EL et al. Characteristics of the sentinel lymph node in breast cancer predict further involvement of higher-echelon nodes in the axilla: a study to evaluate the need for complete axillary lymph node dissection. *Arch Surg* 2001;136:688-692.

[38] Kim T, Giuliano AE, Lyman GH. Lymphatic mapping and sentinel lymph node biopsy in early-stage breast carcinoma: a metaanalysis. *Cancer* 2006;106:4-16.

[39] Nos C, Harding-MacKean C, Freneaux P et al. Prediction of tumour involvement in remaining axillary lymph nodes when the sentinel node in a woman with breast cancer contains metastases. *Br J Surg* 2003;90:1354-1360.

[40] Barranger E, Coutant C, Flahault A et al. An axilla scoring system to predict non-sentinel lymph node status in breast cancer patients with sentinel lymph node involvement. *Breast Cancer Res Treat* 2005;91:113-119.

[41] Degnim AC, Reynolds C, Pantvaitya G et al. Nonsentinel node metastasis in breast cancer patients: assessment of an existing and a new predictive nomogram. *Am J Surg* 2005;190:543-550.

[42] Van Zee KJ, Manasseh DM, Bevilacqua JL et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol* 2003;10:1140-1151.

[43] Viale G, Maiorano E, Pruneri G et al. Predicting the risk for additional axillary metastases in patients with breast carcinoma and positive sentinel lymph node biopsy. *Ann Surg* 2005;241:319-325.

[44] Kohrt HE, Olshen RA, Bermas HR et al. New models and online calculator for predicting non-sentinel lymph node status in sentinel lymph node positive breast cancer patients. *BMC Cancer* 2008;8:66.

[45] Pal A, Provenzano E, Duffy SW et al. A model for predicting non-sentinel lymph node

- metastatic disease when the sentinel lymph node is positive. *Br J Surg* 2008;95:302-309.
- [46] Saidi RF, Dudrick PS, Remine SG, Mittal VK. Nonsentinel lymph node status after positive sentinel lymph node biopsy in early breast cancer. *Am Surg* 2004;70:101-105; discussion 105.
- [47] Alran S, De Rycke Y, Fourchette V et al. Validation and limitations of use of a breast cancer nomogram predicting the likelihood of non-sentinel node involvement after positive sentinel node biopsy. *Ann Surg Oncol* 2007;14:2195-2201.
- [48] Ponzzone R, Maggiorotto F, Mariani L et al. Comparison of two models for the prediction of nonsentinel node metastases in breast cancer. *Am J Surg* 2007;193:686-692.
- [49] Dauphine CE, Haukoos JS, Vargas MP et al. Evaluation of three scoring systems predicting non sentinel node metastasis in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol* 2007;14:1014-1019.
- [50] Chagpar AB, Scoggins CR, Martin RC, 2nd et al. Predicting patients at low probability of requiring postmastectomy radiation therapy. *Ann Surg Oncol* 2007;14:670-677.
- [51] Katz A, Smith BL, Golshan M et al. Nomogram for the prediction of having four or more involved nodes for sentinel lymph node-positive breast cancer. *J Clin Oncol* 2008;26:2093-2098.
- [52] Recht A, Edge SB, Solin LJ et al. Postmastectomy radiotherapy: clinical practice guidelines of the American Society of Clinical Oncology. *J Clin Oncol* 2001;19:1539-1569.
- [53] Harris JR, Halpin-Murphy P, McNeese M et al. Consensus Statement on postmastectomy radiation therapy. *Int J Radiat Oncol Biol Phys* 1999;44:989-990.
- [54] Campbell HE, Taylor MA, Harris AL, Gray AM. An investigation into the performance of the ADJUVANT! Online prognostic programme in early breast cancer for a cohort of patients in the UK. *J Clin Oncol* 2009;in press.
- [55] Bonnefoi H, Potti A, Delorenzi M et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. *Lancet Oncol* 2007;8:1071-1078.
- [56] Buyse M, Loi S, van't Veer L et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006;98:1183-1192.
- [57] Loi S, Haibe-Kains B, Desmedt C et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 2007;25:1239-1246.
- [58] Ma XJ, Wang Z, Ryan PD et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* 2004;5:607-616.
- [59] Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817-2826.
- [60] Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-752.
- [61] Potti A, Dressman HK, Bild A et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006;12:1294-1300.
- [62] Potti A, Mukherjee S, Petersen R et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006;355:570-580.
- [63] Sorlie T, Perou CM, Tibshirani R et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;98:10869-10874.
- [64] Sorlie T, Tibshirani R, Parker J et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003;100:8418-8423.
- [65] Sotiriou C, Neo SY, McShane LM et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA* 2003;100:10393-10398.
- [66] Sotiriou C, Wirapati P, Loi S et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006;98:262-272.
- [67] Van de Vijver MJ, He YD, van't Veer LJ et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.
- [68] Van't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-536.

- [69] Stephenson AJ, Smith A, Kattan MW et al. Integration of gene expression profiling and clinical variables to predict prostate carcinoma recurrence after radical prostatectomy. *Cancer* 2005;104:290-298.
- [70] Patel S, Macdonald OK, Suntharalingam M. Evaluation of the use of prophylactic cranial irradiation in small cell lung cancer. *Cancer* 2009;115:842-850.
- [71] Slotman B, Faivre-Finn C, Kramer G et al. Prophylactic cranial irradiation in extensive small-cell lung cancer. *N Engl J Med* 2007; 357:664-672.
- [72] Stuschke M, Eberhardt W, Pottgen C et al. Prophylactic cranial irradiation in locally advanced non-small-cell lung cancer after multimodality treatment: long-term follow-up and investigations of late neuro-psychologic effects. *J Clin Oncol* 1999;17: 2700-2709.
- [73] Noel G. [Brain metastasis radiotherapy]. *Cancer Radiother* 2006;10:437-443.