

## Apport de l'intelligence artificielle aux données multi-omiques dans les cancers du sein traités par chimiothérapie néo-adjuvante

Laurent Gorvel<sup>1</sup>, Anne-Sophie Chretien<sup>1,3</sup>, Stéphane Fattori<sup>1</sup>, Marie-Sarah Rouviere<sup>1</sup>, Philippe Rochigneux<sup>2</sup>, Anthony Goncalves<sup>2,3</sup>, Daniel Olive<sup>1,3</sup>

<sup>1</sup>Équipe Immunologie des tumeurs, plateforme d'immunomonitoring, centre de recherche en cancérologie de Marseille, Marseille, France.

<sup>2</sup>Département d'oncologie médicale, Institut Paoli-Calmettes, Marseille, France.

<sup>3</sup>Aix-Marseille université, France. daniel.olive@inserm.fr

► L'arsenal thérapeutique dans le cancer du sein s'est enrichi au cours de la dernière décennie. Néanmoins, l'une des questions les plus prégnantes dans l'arbre décisionnel du traitement de ce cancer reste l'identification des facteurs prédictifs de l'efficacité d'un traitement ciblé, à l'image de l'utilisation de l'anticorps monoclonal trastuzumab, dirigé contre le récepteur 2 du facteur de croissance épidermique (HER2), dans le cas de la surexpression de ce récepteur par les cellules tumorales (tumeur HER2<sup>+</sup>). L'efficacité du traitement d'une tumeur dépend certes de la nature de la tumeur et des perturbations qu'elle provoque localement et à distance, mais aussi de facteurs intrinsèques liés à l'hôte, notamment de la réponse de son système immunitaire. L'étude de Sammut *et al.*, que nous analysons ici, concerne l'évaluation de critères prédictifs de l'efficacité de la chimiothérapie néo-adjuvante des cancers du sein, un traitement proposé pour réduire la taille de la tumeur et permettre une chirurgie mammaire conservatrice. L'étude porte sur l'écosystème complexe « tumeur, stroma, vascularisation, et système immunitaire ». Elle complète l'étude publiée en janvier 2022, portant sur 5 161 patientes, qui a montré que le score RCB (*residual cancer burden*) évalué lors de l'examen anatomo-pathologique de la pièce opératoire, après la chimiothérapie néo-adjuvante, est un facteur pronostic dans l'ensemble

des cancers du sein [1]. Le statut RCB est donc un élément essentiel du suivi des patientes bénéficiant du traitement néo-adjuvant. On distingue des réponses complètes (pCR [*complete pathologic response*], RCB-0) et des réponses partielles, que l'on peut classer en trois groupes, allant de la présence d'un résidu tumoral minimal (RCB-I) à un résidu important (RCB-III).

Dans l'étude de Sammut *et al.*, les traitements néo-adjuvants ont été réalisés chez 180 femmes atteintes de différentes formes de cancer du sein, correspondant aux principaux groupes histologiques. Le traitement reposait sur une chimiothérapie associant taxane et anthracycline. Les patientes dont les tumeurs étaient HER2<sup>+</sup> ont reçu trois cycles de thérapie anti-HER2 en association avec un taxane. Au terme du traitement néo-adjuvant, la réponse thérapeutique est classifiée en fonction du score RCB. L'analyse combine les paramètres cliniques, génomiques et transcriptomiques, ainsi que l'étude de l'infiltration de la tumeur par les cellules immunitaires. Les variables associées à la réponse clinique sont ensuite filtrées (sélection univariée et réduction de colinéarité)<sup>1</sup>, puis utilisées pour construire un classifieur utilisant trois

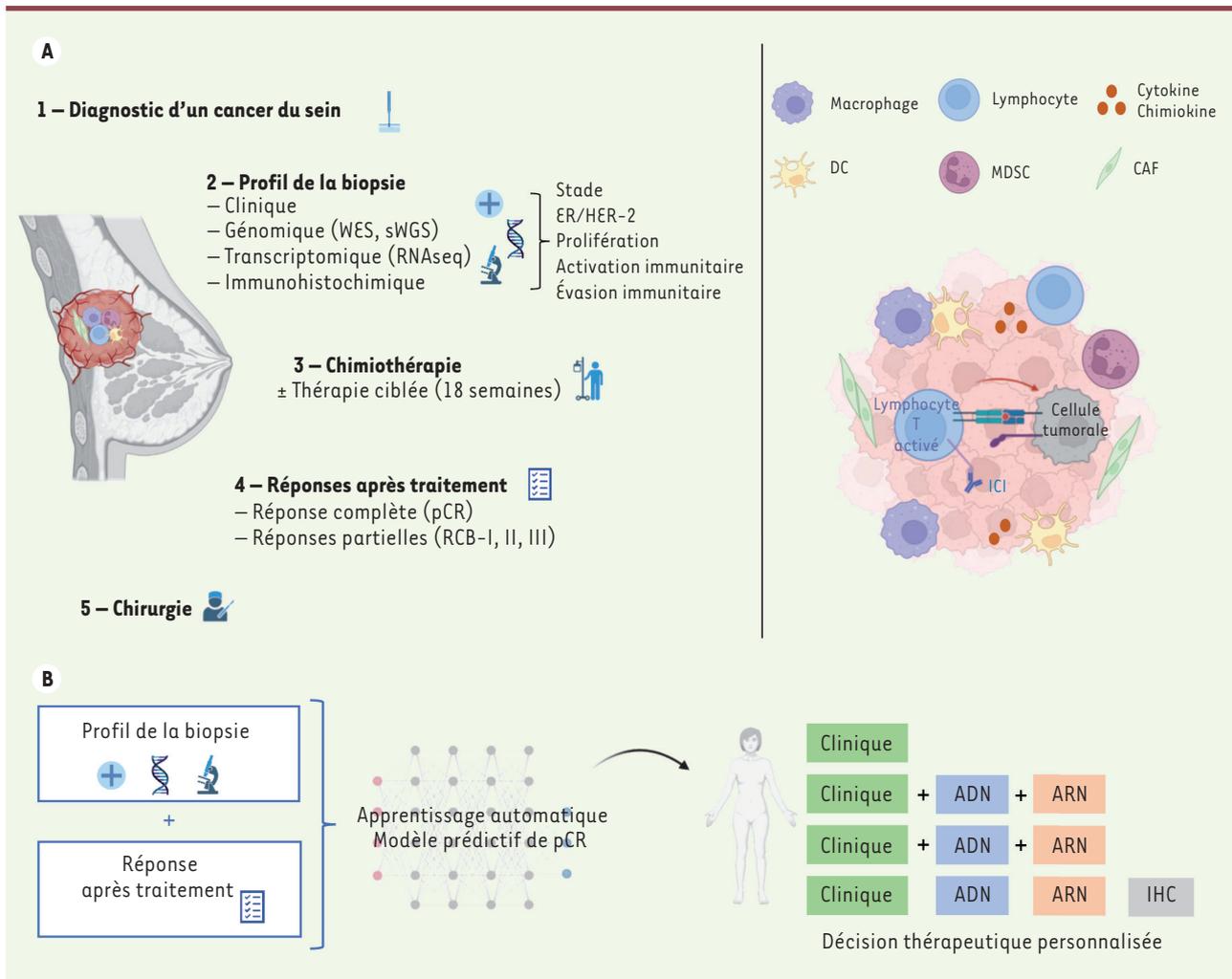
algorithmes : régression logistique, support vecteur machine et forêt aléatoire<sup>2</sup>. Cette approche a mis en évidence les facteurs pertinents, dans le cadre d'un groupe de validation, pour prédire la réponse de la tumeur au traitement. Les auteurs ont décrit un ensemble de critères moléculaires associés à l'existence d'une réponse complète (pCR) chez les patientes HER2<sup>-</sup> : l'absence d'expression du récepteur des œstrogènes (ER) par les cellules tumorales (tumeur ER<sup>-</sup>), les mutations du gène codant le facteur de transcription TP53 (*tumor protein 53*) impliqué dans le contrôle du cycle cellulaire, le nombre de néoantigènes prédits et la charge mutationnelle. De plus, un déficit des mécanismes de recombinaison homologue et une signature mutationnelle APOBEC (*apolipoprotein B mRNA-editing catalytic polypeptide-like*, souvent associée aux cancers du sein HER2<sup>-</sup> avec expression des récepteurs des œstrogènes ou de la progestérone et à une bonne réponse à l'immunothérapie), combinés avec une augmentation de l'instabilité chromosomique, ainsi que le degré de prolifération des cellules tumorales,

<sup>1</sup> La sélection univariée permet d'identifier les variables associées au pronostic. La réduction de colinéarité permet de limiter le poids des variables fortement corrélées entre elles, qui apporteraient des informations redondantes pour le modèle.

<sup>2</sup> Régression logistique : modèle statistique permettant d'analyser les relations entre des variables qualitatives ou quantitatives et une variable qualitative.

Support vecteur machine (ou séparateur à vastes marges) : méthode d'apprentissage supervisé permettant l'identification d'une frontière de séparation (marge) maximale entre les individus de différents groupes.

Forêt aléatoire ou forêt d'arbres décisionnels : méthode d'apprentissage supervisé permettant de construire des arbres de décision servant de modèle prédictif.



**Figure 1. Données multi-omiques dans les cancers du sein traités par chimiothérapie néo-adjuvante.** **A.** À gauche : schéma de prise en charge d'une patiente dans le cadre d'un protocole de chimiothérapie néo-adjuvante. À droite : représentation du microenvironnement immunitaire de la tumeur et représentation de l'activation d'un lymphocyte T par un médicament inhibiteur de point de contrôle immunitaire (ICI). **B.** Intégration des données dans des modèles d'apprentissage automatique, afin de définir une approche thérapeutique personnalisée et de prédire la réponse à ce traitement. DC : *dendritic cell*, cellule dendritique ; MDSC : *myeloid-derived suppressor cell*, cellule suppressive dérivée d'une cellule myéloïde ; CAF : *cancer-associated fibroblast*, fibroblaste associé au cancer ; WES : *whole exome sequencing*, séquençage d'exome complet ; sWGS : *shallow whole genome sequencing*, séquençage du génome complet (moins résolutif) ; RNAseq : *RNA sequencing*, séquençage d'ARN.

évalué par l'analyse transcriptomique, sont également associés à la pCR. Certains de ces paramètres étaient spécifiquement associés à la pCR des tumeurs HER2<sup>-</sup> (charge mutationnelle, néoantigènes et déficit des mécanismes de recombinaison homologue).

L'analyse des paramètres du système immunitaire s'est fondée sur la génomique, la transcriptomique et l'anatomopathologie digitale. Les pertes d'hétérozygotie au locus du complexe majeur d'histocompatibilité de classe

I (CMH I) étaient présentes dans 29 cas et pourraient affecter des allèles impliqués dans la présentation des néoantigènes tumoraux au système immunitaire. Elles sont associées à une moins bonne réponse de la tumeur à la chimiothérapie néo-adjuvante, ce qui suggère un processus d'échappement au système immunitaire. Le résultat de l'analyse transcriptomique indique que la pCR est également associée au degré d'activation du système immunitaire.

Le rôle du microenvironnement immunitaire est confirmé par l'analyse anatomopathologique digitale des biopsies colorées par hématoxyline-éosine, qui met en évidence le caractère prédictif favorable de l'infiltration lymphocytaire ainsi que des marqueurs de cytotoxicité issus d'une signature transcriptomique (score d'activité cytolytique). De plus, l'abondance des acteurs de l'immunité innée (lymphocytes NK [*natural killer*]) et acquise (lymphocytes T CD8<sup>+</sup>, Th1, et B) avant la chimiothérapie est associée

à la pCR, chez les patientes ER<sup>+</sup>HER2<sup>-</sup> comme chez les patientes HER2<sup>+</sup>. En revanche, la présence des mastocytes est associée à un score RCB plus élevé. Cependant, 25 des 46 patientes ayant à la fois une signature d'activation du système immunitaire et de prolifération tumorale n'ont pas atteint la pCR. Elles présentaient des signes évoquant un dysfonctionnement du système immunitaire, qui a été étudié plus en détail chez quatre d'entre elles. L'analyse différentielle des gènes a mis en évidence des signatures transcriptomiques caractéristiques d'un dysfonctionnement des lymphocytes T et d'un enrichissement en cellules immunosuppressives, telles que les lymphocytes T régulateurs, en macrophages de type M2, et en fibroblastes.

À l'issue de ces analyses univariées, les auteurs ont intégré les paramètres anatomopathologiques, génomiques et transcriptomiques associés à la réponse clinique par une approche fondée sur l'intelligence artificielle. La pertinence des résultats a, de plus, été testée sur un « groupe de validation » composé de 75 patientes (18 HER2<sup>+</sup>, 9ER<sup>-</sup>HER2<sup>-</sup>, et 48ER<sup>+</sup>HER2<sup>-</sup>), pour lesquelles des données de séquençage de l'ADN et de l'ARN ainsi que des données numériques d'anatomopathologie étaient disponibles. Les critères associés à la pCR sont l'âge, la présence d'un infiltrat lymphocytaire dans la tumeur, l'expression des récepteurs de la progestérone et des œstrogènes par les cellules tumorales ainsi qu'une plus forte expression de HER2.

Cette étude portant sur le cancer du sein, qui utilise l'intelligence artificielle et prend en compte non seulement la prolifération des cellules tumorales mais aussi l'écosystème tumoral, permet de comparer leurs importances respectives dans la réponse de la tumeur à la chimiothérapie néo-adjuvante. Pour autant, cette étude comporte certaines limites. Tout d'abord, l'utilisation d'un anticorps monoclonal dirigé contre HER2 chez les patientes dont les cellules tumorales expriment ce récepteur implique la mise en œuvre d'une cyto-

toxicité à médiation cellulaire et d'une phagocytose dépendante des anticorps, qui ne dépendent ni de la prolifération tumorale, ni de l'activation du système immunitaire, mais qui dépendent en revanche des niveaux d'expression, par les monocytes et macrophages, ainsi que par les lymphocytes NK et les lymphocytes T $\gamma\delta$ , des récepteurs de la région Fc des Ig(immunoglobulines)G (Rf $\gamma$ ) activateurs et inhibiteurs. Un élément important est la localisation tissulaire de ces cellules immunitaires [2]. On connaissait depuis de nombreuses années son importance dans les cancers de l'ovaire et du côlon, mais les études destinées à évaluer le rôle de ce microenvironnement dans l'efficacité des chimiothérapies néo-adjuvantes des cancers du sein sont plus récentes [3-5]. Leurs résultats sont en faveur d'un terrain favorable, créé par le système immunitaire, permettant une réponse optimale à la chimiothérapie ou encore aux radiations ionisantes [6]. Une autre limite de cette étude concerne la recherche de perte d'hétérozygotie des gènes codant les molécules du CMH I dans les cellules tumorales et la recherche de mutations dans les gènes codant des néoantigènes de ces cellules, potentiellement cibles de lymphocytes T. Ces mutations jouent un rôle important : elles pourraient être à l'origine d'un processus de présentation de ces néoantigènes mutés, en association avec des molécules du CMH I, par des cellules présentatrices d'antigènes, entraînant la sélection d'effecteurs immunitaires dirigés contre les cellules tumorales, mais elle pourraient aussi être à l'origine d'un échappement à ces lymphocytes T par l'apparition de nouveaux mutants. En effet, perte d'hétérozygotie des gènes codant les molécules du CMH I et mutations des néoantigènes pourraient limiter leur présentation et leur reconnaissance par des lymphocytes T spécifiques, donc l'efficacité de la réponse immunitaire anti-tumorale. Il aurait été pertinent de chercher à savoir si, au sein de ce groupe de patientes,

d'autres modifications, affectant en particulier l'immunité innée, pourraient compenser ce déficit fonctionnel des lymphocytes T. Si les paramètres immunologiques associés à la réponse sont déjà connus, en particulier les lymphocytes T CD8<sup>+</sup> et les lymphocytes B, leur étude, uniquement transcriptomique, ne renseigne pas sur leur présence et leur distribution, par exemple dans les structures lymphoïdes tertiaires, sites intra- et péri-tumoraux associés à une meilleure survie [7] et une meilleure réponse au traitement. Cependant, les mécanismes suppresseurs identifiés n'incluent pas les sous-populations de lymphocytes T régulateurs les plus pathogènes, ni les différentes populations de macrophages ; la notion de macrophage M1, pro-inflammatoire, et M2, anti-inflammatoire, ayant récemment évolué au profit d'une classification fondée sur l'appartenance au microenvironnement tumoral (macrophages associés à la tumeur) et sur le recrutement de macrophages dérivés de monocytes du sang. Il convient de noter que les macrophages associés aux tumeurs sont également composés de divers sous-types cellulaires plus ou moins inflammatoires [8]. Il en est de même pour les fibroblastes, actuellement sous-stratifiés pour identifier ceux dotés de l'effet immunosuppresseur maximum [9]. Enfin, l'identification de lymphocytes NK CD56<sup>dim</sup> est surprenante car ceux-ci correspondent en général au stade de différenciation correspondant à des capacités fonctionnelles cytotoxiques. Il conviendrait de réanalyser les données transcriptomiques pour préciser si des anomalies du profil de différenciation apporté par les facteurs de transcription Tbet (*T-cell-specific T-box transcription factor*) et EOMES (*eomesodermin*) ainsi que leur signature d'activité cytotolytique sont détectables. La notion d'exclusion immune étudiée par étude transcriptomique suggère son rôle dans l'absence de pCR. Malheureusement, l'exclusion immune reste peu développée dans l'article.



En conclusion, cette étude est une des premières à introduire les concepts d'apprentissage automatique apportés par l'intelligence artificielle pour nous aider à extraire les marqueurs prédictifs les plus importants de l'efficacité de la chimiothérapie néo-adjuvante des cancers du sein. Malgré la performance de ces modèles d'apprentissage automatique pour la réponse au traitement néo-adjuvant par rapport à ceux fondés sur des variables cliniques, ce classifieur reste imparfait et l'intégration d'autres paramètres, tels que ceux fournis par le transcriptome ou l'imagerie sur coupes de la tumeur, devrait permettre d'améliorer la prédiction de la réponse au traitement et d'alimenter l'arbre décisionnel. Les auteurs font un parallélisme avec les études actuelles identifiant des marqueurs immunitaires prédictifs de la réponse aux immunothérapies par anticorps monoclonaux dirigés contre les molécules de cosignalisation (inhibiteurs de point de contrôle immunitaire), qui sont des molécules induisant un signal activateur ou inhibiteur, complétant le signal provoqué par la présentation d'un antigène à la synapse immunologique (par exemple, CTLA-4

ou PD-1). Les études les plus récentes indiquent le rôle des populations de lymphocytes CD4<sup>+</sup> de profil Th1 [10]. Par ailleurs, elles indiquent que deux sous-populations de cellules immunitaires seraient induites : celle des lymphocytes T dits « épuisés » (*exhausted T cells*), dont certains produisent le granzyme K, une molécule cytotoxique leur permettant de conserver une activité anti-tumorale, et celle issue de cellules des organes lymphoïdes et du sang, possédant globalement une meilleure activité anti-tumorale, qui semblent marquer une réponse au traitement par chimiothérapie néo-adjuvante [11, 12]. Il conviendra de préciser le rôle de ces sous-populations dans la réponse aux traitements néo-adjuvants. ♦

### Artificial intelligence contribution to multi-omic data in the treatment of breast cancers by neoadjuvant chemotherapy

#### LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

#### RÉFÉRENCES

1. Yau C, Osdoit M, van der Noordaa M, et al. Residual cancer burden after neoadjuvant chemotherapy and

long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5 161 patients. *Lancet Oncol* 2022 ; 23 : 149-60.

2. Kuettl L, Catena R, Ozcan A, et al. Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. *Nat Cancer* 2022 ; 3 : 122-33.
3. Ali HR, Dariush A, Provenzano E, et al. Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res* 2016 ; 18 : 21.
4. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006 ; 313 : 1960-4.
5. Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 2003 ; 348 : 203-13.
6. Green DR, Ferguson T, Zitvogel L, Kroemer G. Immunogenic and tolerogenic cell death. *Nat Rev Immunol* 2009 ; 9 : 353-63.
7. Dieu-Nosjean MC, Antoine M, Danel C, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J Clin Oncol* 2008 ; 26 : 4410-7.
8. Binnewies M, Pollack JL, Rudolph J, et al. Targeting TREM2 on tumor-associated macrophages enhances immunotherapy. *Cell Rep* 2021 ; 37 : 109844.
9. Huang TX, Tan XY, Huang HS, et al. Targeting cancer-associated fibroblast-secreted WNT2 restores dendritic cell-mediated antitumor immunity. *Gut* 2022 ; 71 : 333-44.
10. Bassez A, Vos H, Van Dyck L, et al. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat Med* 2021 ; 27 : 820-32.
11. Nagasaki J, Inozume T, Sax N, et al. PD-1 blockade therapy promotes infiltration of tumor-attacking exhausted T cell clonotypes. *Cell Rep* 2022 ; 38 : 110331.
12. Liu B, Hu X, Feng K, et al. Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. *Nat Cancer* 2022 ; 3 : 108-21.



Avec m/s, vivez en direct  
les progrès et débats  
de la biologie et de la médecine

CHAQUE MOIS / AVEC LES ARTICLES DE RÉFÉRENCE DE M/S  
CHAQUE JOUR / SUR WWW.MEDECINESCIENCES.ORG

Abonnez-vous sur

[www.medecinesciences.org](http://www.medecinesciences.org)