



HAUTE AUTORITÉ DE SANTÉ

RAPPORT D'ÉVALUATION TECHNOLOGIQUE

Utilité clinique des signatures génomiques dans le cancer du sein de stade précoce

Janvier 2019

Ce rapport d'évaluation technologique réalisé en vue d'une prise en charge par l'assurance maladie obligatoire, est téléchargeable sur www.has-sante.fr

Haute Autorité de santé

Service communication - information

5, avenue du Stade de France – F 93218 Saint-Denis La Plaine Cedex

Tél. : +33 (0)1 55 93 70 00 – Fax : +33 (0)1 55 93 74 00

Sommaire

Acronymes.....	4
Résumé	6
Introduction.....	11
1. Contexte médico-scientifique et réglementaire.....	12
1.1 Utilité clinique d'un nouveau marqueur tumoral.....	12
1.2 Caractéristique techniques des SG en évaluation.....	14
1.3 Population ciblée préalablement pour l'évaluation.....	14
1.4 Évènements cliniques en contexte adjuvant.....	20
1.5 Remontées d'utilisation en vie réelle (données RIHN 2016-2017).....	20
1.6 État des lieux des principales évaluations technologiques nationales.....	21
1.7 Essais cliniques en cours.....	22
2. Protocole de l'évaluation.....	23
2.1 Objectifs de l'évaluation.....	23
2.2 Méthode de l'évaluation.....	23
2.3 Champ d'évaluation.....	23
2.4 Stratégie de recherche et de sélection des publications.....	24
2.5 Critères d'évaluation de l'utilité clinique.....	25
2.6 Groupe de travail multidisciplinaire.....	30
2.7 Recueil du point de vue des parties prenantes.....	32
2.8 Consultation de l'Institut national du cancer.....	32
3. Résultats de l'évaluation.....	33
3.1 Diagramme de flux de la recherche systématique.....	33
3.2 Analyse des données de la littérature.....	34
3.3 Position du groupe de travail.....	43
3.4 Position institutionnelle de l'INCa.....	48
3.5 Points de vue des parties prenantes.....	48
4. Conclusions et perspectives.....	53
4.1 Conclusions.....	53
4.2 Perspectives.....	55
Références bibliographiques.....	57
Fiche descriptive.....	62

Acronymes

ADEL	Association des épidémiologistes de langue française
AHRQ	<i>Agency for Healthcare Research and Quality's</i> (Etats-Unis)
ALD	Affection longue durée
AOL	Adjuvant on line !
ARS	Agence régionale de santé
C (haut/bas)	Risque clinique élevé/faible
CADTH	<i>Canadian Agency for Drugs and Technologies in Health</i> (Canada)
CCAM	Classification commune des actes médicaux
CCPT	Critères clinico-pathologiques traditionnels
CNP	Conseil national professionnel
CTA	Chimiothérapie adjuvante
CTCAE	<i>Common Terminology Criteria for Adverse Events</i>
DATECAN	<i>Definition for the Assessment of Time-to-event Endpoints in CANcer trials</i>
EBCTG	<i>Early Breast Cancer Trialists' Collaborative Group</i>
EGAPP	<i>Evaluation of Genomic Applications in Practice and Prevention</i>
EI	Évènement indésirable
EIG	Évènement indésirable grave
EMA	<i>European medicine agency</i>
ESMO	<i>European Society for Medical Oncology</i>
EUNETHTA	<i>EUropean Network for HTA</i>
FDA	<i>Food and drug administration</i>
G (haut/bas)	Risque génomique élevé/faible
GFCO	Groupe francophone de cytogénomique oncologique
GRADE	<i>Grading of Recommendations Assessment, Development and Evaluation Group</i>
HAS	Haute autorité de santé
HER2	<i>Human Epidermal Growth Factor Receptor-2</i>
HR	<i>Hazard Ratio</i>
HTA	Hormonothérapie adjuvante
IC	Intervalle de confiance (à 95 %)
ICHOM	<i>International Consortium for Health Outcomes Measurement</i>
IQWiG	<i>Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen</i>
INCa	Institut national du cancer
INESS	Institut National d'Excellence en Santé et en services sociaux (Québec)
ITT	Intention de traiter
KCE	Centre fédéral d'expertise des soins de santé (Belgique)
MSAC	<i>Medical Services Advisory Committee</i> (Australie)
NABM	Nomenclature des actes de biologie médicale

NCCN	<i>National Comprehensive Cancer Network</i>
NCI	<i>National Cancer Institut</i>
NCT	<i>National Clinical Trials</i> (registre nord-américain des essais cliniques)
NHS	<i>National Health Service</i> (Royaume-Uni)
NSN	Nombre de sujets nécessaires
PP	<i>Per protocole</i>
RCP	Réunion de concertation pluridisciplinaire
RIHN	Référentiel des actes innovants hors nomenclature
RIHN	<i>Risk Of Bias In Non-randomized Studies - of Interventions</i>
SBR	<i>Scarff-Bloom et Richardson</i>
SFCO	Société française de chirurgie oncologique
SFMPP	Société française de médecine prédictive et personnalisée
SFSPM	Société française de sénologie et des pathologies mammaires
SSM	Survie sans maladie
SSMD	Survie sans métastase à distance
SSMI	Survie sans maladie invasive
STEEP	<i>Standardized Definitions for Efficacy End Point in Adjuvant Breast Cancer Trial</i>
SWOG	<i>Southwest Oncology Group</i>
RCP	Réunion de concertation pluridisciplinaire
RO	Récepteurs aux œstrogènes
RH	Récepteurs hormonaux
ROR	<i>Risk of recurrence (Prosigna)</i>
RP	Récepteurs à la progestérone
RS	<i>Recurrence score</i> (Oncotype Dx)
SG	Signature génomique
TNM	<i>Tumor, node, metastasis</i>

Résumé

Objectifs

L'objectif de cette autosaisine de la Haute autorité de santé (HAS) est :

- d'évaluer chez des patientes atteintes d'un cancer infiltrant du sein de stade précoce susceptibles de recevoir en pratique une chimiothérapie adjuvante (CTA), l'utilité clinique de l'utilisation d'une signature génomique (SG) pour optimiser cette décision thérapeutique ;
- d'évaluer la concordance des décisions entre les quatre SG¹ inscrites actuellement aux Référentiel des actes innovants hors nomenclatures (RIHN) ;
- d'émettre un avis sur le bienfondé de leur remboursement par l'Assurance Maladie.

Contexte de l'évaluation

Une CTA est administrée dans le but de réduire le risque de récurrence de la maladie après ablation de la tumeur et d'allonger la survie des patientes. Cette décision prend en compte le degré d'extension et le profil biologique de la tumeur ainsi que l'état général et les préférences de la patiente. Pour une part des patientes atteintes d'un cancer du sein de stade précoce, la CTA n'est pas systématique en raison d'un bénéfice attendu relativement modeste lié à un pronostic naturellement favorable. En cas de risque intermédiaire de récurrence au vu des critères clinico-pathologiques traditionnels (CCPT)², cette décision de CTA resterait dans certains cas difficile et suffisamment hétérogène entre les centres pour en faire un sujet de préoccupation médicale.

La population préalablement ciblée pour l'évaluation des SG correspond aux patientes à risque intermédiaire qui pourraient être « au cas par cas » plus particulièrement susceptibles³ de recevoir une CTA en complément d'une hormonothérapie adjuvante (HTA). Ces femmes possèdent une tumeur infiltrante de stade pT1c-T2⁴ pN0-N1⁵, hormono-dépendante, de statut immunohistochemique RO+⁶/HER2- (environ 40 % des cancers infiltrants).

En contexte français, cette population préalablement ciblée est constituée en réalité de trois sous-groupes de patientes :

- celui pour qui il n'existe pas d'indication préalable de CTA⁷ ;
- celui pour qui l'indication de CTA reste incertaine⁸ ;
- celui pour qui une indication⁹ de CTA est préalablement établie.

L'ajout d'une SG en supplément des CCPT pourrait améliorer la pertinence de la décision thérapeutique soit (i) en levant l'incertitude chez celles dont l'indication de CTA reste équivoque, (ii) en désescaladant le traitement chez celles dont l'indication préalable de CTA est pourtant établie. Dans l'idéal, les patientes pourraient, sans perte de chances, ne plus se voir proposer ce traitement générateur de complications (si score génomique faible), celui-ci restant réservé à celles pour

¹ Oncotype Dx, MammaPrint, Prosigna et Endopredict (Acte RIHN : N537)

² Âge de la patiente, taille tumorale, envahissement ganglionnaire, type histologique, grade histopronostique, statut des récepteurs hormonaux et HER2, embolies vasculaires

³ Les tumeurs infracentimétriques (pT1ab) RO+/HER2- n'ont pas été retenues pour cette évaluation car leur pronostic est globalement excellent sous HTA exclusive ; le recours à une CTA n'étant décidé en pratique que dans des cas complexes et peu fréquents

⁴ Taille tumorale comprise entre un et cinq cm.

⁵ Sans invasion ganglionnaire (pN0) ou avec une invasion limitée à un à trois ganglions axillaires (pN1)

⁶ Récepteurs aux œstrogènes (RO) ± progestérone (RP)

⁷ Notamment, la plupart des tumeurs RO+/HER2- de stade précoce sans invasion ganglionnaire (pN0) et de grade 1 histopronostique

⁸ Les CCPT pouvaient être équivoques ou bien discordants

⁹ Dans le cadre de la population de cette évaluation, la présence d'au moins deux critères de mauvais pronostic ou d'un grade 3 ou d'une extension ganglionnaire macrométastatique (pN1) suffisait pour poser une indication préalable.

qui le bénéficiaire serait suffisant (si score génomique élevé). En décembre 2013, le dernier rapport de l'Institut National du Cancer (INCa) n'avait pu conclure en faveur d'une démonstration de la valeur ajoutée et de l'utilité clinique des SG de première génération par rapport aux CCPT (Oncotype Dx et Mammaprint).

Méthode

Dans un premier temps, un rapport provisoire a été réalisé. Il s'est appuyé sur :

- une revue systématique de la littérature à la recherche des études prospectives et comparatives d'impact clinique¹⁰ et de concordance décisionnelle entre les différentes SG ;
- une analyse critique des données factuelles ;
- la consultation d'experts extérieurs réunis au sein d'un groupe de travail (GT) multidisciplinaire¹¹.

Dans un second temps, la HAS a recueilli lors d'une phase de relecture du rapport provisoire :

- la position institutionnelle de l'Institut National du Cancer (INCa) ;
- le point de vue des parties prenantes¹² concernant sa clarté et sa lisibilité, leur appropriation du travail et les conséquences pratiques et organisationnelles qu'ils souhaitent faire remonter.

Résultats

Quatre études ont pu être sélectionnées. Aucune d'entre elles n'a permis de hiérarchiser les différentes SG sur la base de leur utilité clinique. Aucune étude d'utilité clinique n'a été identifiée concernant les tests de deuxième génération (Prosigna et Endopredict). Aucune étude d'utilité clinique n'a été identifiée concernant la valeur ajoutée d'une SG au sein d'un algorithme décisionnel intégrant préalablement les CCPT (modèle de prédiction du risque).

Toutefois, l'analyse des quatre études retenues a permis, en accord avec la position des experts du GT :

- de faire ressortir la complexité des publications scientifiques sélectionnées ;
- de montrer que les SG ne peuvent se substituer aux CCPT pour optimiser la décision de CTA ;
- de montrer, qu'à ce jour, les SG ne peuvent être utilisées comme tests prédictifs de l'efficacité d'un protocole de CTA ;
- de montrer que la décision de CTA pouvait être divergente dans 20 à 25 % des cas entre SG pour un même échantillon de patientes (étude OPTIMA Prelim) ;
- de montrer, qu'à ce jour, il n'existe pas de preuves directes en faveur de l'utilité clinique des SG de première génération en supplément des CCPT (Oncotype Dx et Mammaprint).

En effet, pour ce dernier point les résultats analysés sont principalement issus de deux essais internationaux MINDACT et TAILORx initiés il y a plus de 11 ans, au moment de l'émergence des SG. De plus, l'un des deux (TAILORx) a été réalisé en contexte nord-américain. Ces éléments expliquent en grande partie le fait que les objectifs et les populations ciblées par ces études étaient globalement différents et volontairement plus larges que la population susceptible de recevoir une CTA en contexte français.

Dans cette population d'intérêt spécifiquement ciblée par la HAS, les deux études n'ont pas rapporté de données probantes ou suffisantes (peu ou pas d'analyses de sous-groupe de patientes en incertitude décisionnelle ou avec une indication préalable de CTA). Dans l'étude TAILORx les seuils de décision thérapeutique d'Oncotype Dx ne pouvaient être valables en raison de risques de biais

¹⁰ L'amélioration de l'état de santé des patientes sous la stratégie de décision basée sur une SG par rapport à la stratégie habituelle basée sur les CCPT. Les études (pro)rétrospectives de validation des performances pronostiques ou prédictives de ces tests (validation clinique) et les études médico-économiques n'ont pas été retenues.

¹¹ Oncologues, biologistes, pathologistes, chirurgien sénologue, méthodologistes et biostatisticiens spécialisés en oncologie, une patiente

¹² Sept organismes professionnels et deux associations de patients

majeurs, de plus la valeur ajoutée pronostique du test n'avait pas été explorée. Aucune donnée de sécurité et de qualité de vie n'a été identifiée dans ces études pour établir la balance bénéfice/sécurité de la nouvelle stratégie décisionnelle basée sur une SG. La grande majorité des organismes professionnels, les deux associations de patientes et l'INCa partageaient les analyses et les propositions des experts du GT concernant le besoin de clarification et d'un encadrement plus restrictif de l'utilisation des SG dans le cancer du sein.

Conclusion

Si selon certains experts du GT, une valeur pronostique intrinsèque des SG est possible, l'ensemble des membres fait état de l'absence ou du manque de données probantes démontrant actuellement leur utilité clinique dans les populations d'intérêt en contexte français. Par ailleurs, eu égard à leur objectif initial de développement et aux données analysées, les quatre SG ne peuvent revendiquer actuellement une valeur prédictive concernant l'efficacité d'une CTA.

À l'issue de ce travail, **la HAS rappelle le rôle essentiel des CCPT** pour la prise de décision de CTA **en contexte français** et **considère que les SG ne devraient pas être utilisées pour s'y substituer**. L'utilité clinique n'a pas été actuellement démontrée concernant les tests de première génération (Oncotype Dx/Mammaprint), faute de données probantes ou suffisantes dans une population préalablement ciblée pour l'évaluation. L'utilité clinique n'a pas été actuellement démontrée concernant les tests de deuxième génération (Prosigna/Endopredict), faute d'étude disponible.

La HAS considère que leur utilisation ne présente **aucun intérêt** chez les patientes :

- sans aucune indication de CTA sur la base des CCPT (patiente de bon pronostic sans aucune indication de CTA) ;
- présentant une tumeur non hormonodépendante (RO-) ou HER2+ ou triple négatifs.

La HAS considère que dans l'attente des résultats des trois essais en cours¹³ dans une population à haut risque clinique (notamment N+), leur utilisation **ne peut être actuellement indiquée** en vue d'une désescalade thérapeutique chez les patientes :

- qui présentent un envahissement ganglionnaire macrométastatique (patientes N+) ;
- qui présentent une indication préalable de CTA sur des critères de mauvais pronostic.

Proposition 1 : Mieux évaluer et contrôler la diffusion des SG

La HAS considère, qu'en l'état actuel des connaissances, la population qui présente potentiellement un intérêt pour évaluer l'utilisation des SG est celle **en incertitude décisionnelle**.

- En conformité avec les recommandations de **Saint Paul de Vence** (2015), de **Saint Gallen** (2017) et la **position du GT**, la HAS propose d'encadrer cette utilisation au niveau d'une sous-population bien définie et plus restreinte de **cas équivoques** ayant un degré de prolifération **intermédiaire** (environ 5 à 10 % des cancers invasifs). Ces patientes devraient présenter un **adénocarcinome infiltrant** (préférentiellement canalaires), **pT1c-2 pN0-N1mi¹⁴** de **grade 2** et **sans indication préalable de CTA au vu des autres critères de mauvais pronostic¹⁵** ;
- toutefois, la HAS n'est pas en mesure de préconiser cette utilisation **en dehors du cadre de la recherche et de l'innovation** du fait de l'absence de preuves directes en faveur d'une utilité clinique ;
- la décision de prescription de SG pour les patientes éligibles devrait être **collégiale** (analyse des CCPT en réunion de concertation pluridisciplinaire) et après **accord de la patiente** (devant une information claire, loyale et appropriée du professionnel et le recueil de ses préférences en matière de choix thérapeutique) ;

¹³ RXPONDER, WSG ADPAT (RO+/HER2-), OPTIMA

¹⁴ Envahissement micrométastatique

¹⁵ Âge jeune et présence d'embolies lymphovasculaires

- la HAS considère qu'il existe un besoin de poursuivre la recherche clinique chez les patientes en incertitude décisionnelle telles que définies plus haut. À cet effet, elle **propose le maintien d'un financement de recherche** conditionné à l'obligation d'un recueil exhaustif et contributif de données cliniques comparatives pour toute SG éligible à ce type de financement (RIHN notamment).
- le **schéma d'étude** nécessaire pour une démonstration formelle de l'utilité clinique d'une SG requerrait un essai prospectif randomisé à quatre bras (double randomisation) ;
- A défaut, la HAS propose un schéma d'étude alternatif, **adaptatif** en deux temps :
 - ▶ **Phase préliminaire observationnelle** (phase 1) estimant la valeur ajoutée « analytique »¹⁶ des valeurs de score génomique obtenues **en insu des CCPT** pour chaque signature dans cette population étroite et bien définie ;
 - ▶ **Phase interventionnelle** (phase 2) : si la valeur ajoutée analytique est suffisante, il serait nécessaire d'attribuer alors exclusivement la CTA sur la base du score génomique obtenu **en insu des CCPT** par rapport au seuil de décision thérapeutique¹⁷ du test. L'estimation de l'intérêt thérapeutique de la CTA guidée par une SG dans cette population serait établie au regard de l'importance clinique de la différence de **survie sans maladie invasive à cinq ans** entre les deux groupes.

La HAS considère par ailleurs que leur utilisation **n'est actuellement pas indiquée** pour les cas complexes et hétérogènes d'incertitude décisionnelle (**discordance clinico-histo-pathologique majeure**). Le recours préférentiel à un modèle de prédiction du risque individuel qui serait à valider en contexte français est un préalable indispensable. Ce type d'algorithme décisionnel pourrait permettre d'estimer la valeur ajoutée d'un score génomique par rapport aux CCPT dans l'optique de l'y intégrer ultérieurement une fois son intérêt démontré.

Proposition 2 : Mieux informer le public et les professionnels concernés

- La HAS considère qu'il est nécessaire de rassurer le public et les patientes sur la problématique française autour de l'administration de la CTA dans le cancer du sein de stade précoce RO+/HER2- qui n'est pas celle du surtraitement de patientes normalement peu susceptibles d'en recevoir ;
- la HAS considère que ses préconisations sur l'utilisation des SG devraient être largement diffusées auprès des professionnels concernés afin de limiter les mésusages éventuels :
 - ▶ risque de sous-traitement lié à une désescalade thérapeutique chez des patientes avec une indication préalable de CTA ;
 - ▶ risque d'escalade thérapeutique non pertinente chez les patientes peu susceptibles de recevoir de la CTA en contexte français.

Perspectives

La HAS considère que le recueil des données dans le cadre du RIHN n'a pas contribué à l'évaluation des SG¹⁸, mais plutôt, et dans la limite de leur exhaustivité, à observer leurs conditions d'utilisation en vie réelle par SG. La HAS préconise **une amélioration du recueil des données cliniques qui devra être exhaustif et contributif en vue d'une évaluation scientifique ultérieure**.

Au-delà de ce travail, la HAS considère comme indispensable que soit élaboré de préférence par l'INCa et conjointement avec les professionnels **un document commun de recommandations nationales de bonne pratique dans le cancer du sein RO+/HER2- de stade précoce** visant à réduire l'hétérogénéité des prescriptions actuelles de CTA entre les centres d'oncologie. Ce travail d'harmonisation concernant les indications et les non indications établies de CTA permettrait de

¹⁶ Etendue de la dispersion/distribution des valeurs de score génomique entre patientes (sans référence à aucun seuil de décision)

¹⁷ Si valeur sous le seuil → HTA exclusive ; si valeur au-dessus du seuil → CTA + HTA

¹⁸ L'étude comparative OPTIGEN n'a pas démarré à ce jour

poursuivre également l'identification des situations d'incertitude décisionnelle (zone grise du clinicien).

Pour résoudre les cas complexes d'incertitude décisionnelle, la HAS considère qu'un algorithme d'aide à la décision intégrant les nombreux CCPT disponibles (au sein d'un modèle multivarié) est devenu indispensable. Certains modèles existent et sont bien avancés¹⁹. Ceux-ci pourraient être adaptés et validés en contexte français ou bien d'autres pourraient être développés *de novo* sur la base de données françaises récentes de bonne qualité.

L'évaluation du **risque de récurrence** observé **sous HTA exclusive** des patientes en incertitude décisionnelle pourrait se faire également à l'aide d'études pronostiques sur des données récentes et de bonne qualité.

Trois essais en cours de réalisation²⁰ sont attendus dans les populations ciblées pour l'évaluation chez qui il existe une indication préalable de CTA (notamment les patientes N+). Il n'existe pas actuellement d'étude d'utilité clinique en cours pour les tests Endopredict et Mammaprint.

¹⁹ Par exemple, le modèle de prédiction PREDICT

²⁰ Oncotype Dx : RxPONDER et WSG ADAPT, Prosigna : OPTIMA

Introduction

Une signature génomique (SG) est un type de biomarqueur²¹ quantifiant, le plus souvent au sein d'une tumeur, l'expression moléculaire d'un panel de gènes sélectionnés (transcriptomes²²). L'objectif de ces tests vise à optimiser une décision médicale. Celle-ci peut concerner le fait d'administrer ou non un traitement en particulier (2).

Le résultat est rendu sous forme d'un **score génomique** reposant sur une construction mathématique complexe (modélisation/algorithmes) qui sera interprété au regard d'un **seuil de décision pré-défini**²³ conduisant la patiente à des prises en charge différentes au même titre que le seuil de positivité d'un test diagnostique (marqueur positif ou négatif) (3).

L'intérêt clinique d'une SG dans le cancer du sein de stade précoce, comme tout marqueur tumoral, peut résider (4, 5) :

- dans sa valeur pronostique (évolution naturelle du cancer en dehors de tout traitement) ;
- dans sa valeur prédictive de l'efficacité anticipée pour une thérapie donnée ;
- dans un mécanisme combinant ces deux propriétés.

L'utilité clinique d'un nouveau marqueur tumoral devrait s'apprécier conjointement (6-9) par :

- un **impact décisionnel** sur les pratiques habituelles : capacité à modifier dans une certaine mesure les décisions médicales ;
- un **impact clinique** sur l'état de santé : existence de *preuves directes*²⁴ de l'amélioration du devenir clinique des patients avec la nouvelle stratégie.

Quatre SG inscrites au **Référentiel des actes innovants hors nomenclatures (RIHN)** sont actuellement financés par un acte global (N137) dans l'indication du cancer du sein de stade précoce pour optimiser la décision de chimiothérapie adjuvante (CTA) : **Oncotype Dx, Mammaprint, Prosigna, Endopredict**. Depuis avril 2016, ce financement est dérogatoire et transitoire avec une obligation de recueil prospectif et comparatif de données cliniques. Il existe une demande forte de clarification scientifique émanant des professionnels de santé et des institutions nationales concernées²⁵ d'évaluer leur utilité clinique. Pour y répondre, la HAS s'est autosaisie en 2016 en vue d'évaluer l'opportunité de leur remboursement par l'Assurance maladie.

²¹ Un biomarqueur est une « caractéristique qui est objectivement mesurée et évaluée comme un indicateur de processus biologiques normaux ou pathologiques, ou de réponses pharmacologiques à une intervention thérapeutique » (1).

²² ARN messager (ARNm) au sein de l'échantillon tumoral

²³ Certaines SG dans le cancer du sein disposent actuellement de deux seuils de décision créant trois catégories de risque génomique

²⁴ Type de preuves permettant d'établir selon le groupe GRADE et de l'EGAPP, un lien de causalité entre l'intervention à évaluer et l'amélioration de la santé des patients par l'intermédiaire d'une étude interventionnelle comparative et randomisée dans une population pertinente et sur des critères cliniques de morbi-mortalité (10, 11)

²⁵ Ministère de la santé, Agences régionales de santé (ARS), Institut national du cancer (INCa)

1. Contexte médico-scientifique et réglementaire

Ce chapitre de contexte a été rédigé à partir d'une revue non systématique de la littérature s'appuyant sur les standards de qualité de la grille PRISMA-P afin d'argumenter le choix des questions d'évaluation (PICOTS), des populations d'intérêt et des critères d'évaluation retenus par la HAS pour cette évaluation (12).

1.1 Utilité clinique d'un nouveau marqueur tumoral

1.1.1 Préambule

Certains tests médicaux sont *diagnostiques* s'ils sont utilisés en cas de suspicion d'une maladie donnée pour aider à discriminer les sujets malades des non malades ; d'autres sont *pronostiques* s'ils sont utilisés en cas de maladie avérée pour aider à discriminer les sujets à risque élevé de complication de ceux à risque plus faible. D'autres encore peuvent être *prédictifs* de d'efficacité anticipée à un traitement²⁶ s'ils sont utilisés pour aider à discriminer les sujets qui répondront suffisamment à une thérapie spécifique (en termes d'efficacité relative)²⁷.

Les étapes de validation d'un test pronostique, sa nature et la manière de mesurer ses performances sont très proches de celles d'un test diagnostique (9, 13-15). Les étapes de validation et la quantification des performances cliniques d'un test prédictif sont différentes et plus complexes que les deux précédents (2, 4, 5, 16-19).

1.1.2 Distinction entre validation et utilité cliniques

La **validation clinique** d'un nouveau test pronostique ou prédictif réside dans sa capacité à montrer respectivement une *différence pronostique* (existence d'un gradient) ou une *différence d'efficacité* (existence d'une interaction test-traitement) statistiquement significative (médecine stratifiée). Pour que la valeur pronostique d'un test soit clairement établie, sa **valeur ajoutée** par rapport aux autres examens réalisés devrait être également *statistiquement indépendante*²⁸ et *cliniquement pertinente* (20, 21). Si un nouveau test coûteux et complexe à réaliser est redondant avec les tests habituels (situation d'absence de valeur ajoutée) alors ce nouveau test est inutile (22).

L'**utilité clinique** d'un nouveau test pronostique ou prédictif est démontrée, si l'on observe un *impact sur les décisions* (condition préalable pour la rendre plausible) produisant un *impact clinique favorable* sur des critères de morbi-mortalité pertinents (7, 23, 24).

1.1.3 Utilité clinique d'une SG pour guider la CTA d'un cancer du sein

L'intérêt pour la santé d'un nouveau marqueur guidant la décision d'un traitement repose sur la démonstration d'un **gain clinique** (bénéfice net suffisant) basée sur une **balance bénéfice sécurité favorable**²⁹. Cette balance bénéfice/sécurité doit être meilleure avec la nouvelle stratégie qu'avec la stratégie habituelle sans le nouveau marqueur (2, 6, 22, 25). L'évaluation de l'utilité d'un nouveau marqueur requière de formaliser d'un côté la stratégie décisionnelle habituelle et de l'autre *sa place et le moment de sa réalisation dans cette stratégie* ainsi que *son rôle et son seuil de décision thérapeutique* (15, 26).

²⁶ Test théranostique ou compagnon d'une thérapie ciblée (4).

²⁷ Réduction du risque relatif (si données binaires) ou *hazard ratio* (si données de survie).

²⁸ Par l'intermédiaire d'un modèle multivarié explicatif (3).

²⁹ L'évaluation de la balance bénéfice sécurité est basée sur une différence entre la réduction de risque absolu de complication liée à la maladie grâce au traitement (bénéfice thérapeutique) et l'augmentation du risque absolu de faire un événement lié à la toxicité du traitement. L'analyse des risques relatifs est moins pertinente et n'est pas adaptée à l'estimation de cette balance (2).

► Efficacité de la CTA dans le cancer localisé du sein

Avec environ 54 000 nouveaux cas pour environ 11 700 décès annuels en France lors de l'année 2013 (27), le cancer du sein est la première cause de cancer féminin. Environ 75 % des nouveaux cas sont infiltrants et ils toucheraient une femme de moins de 75 ans dans 80 % des cas (27, 28).

Une CTA est indiquée pour certains cancers localisés et infiltrants du sein. Son objectif est de diminuer le risque de récurrence tumorale afin d'augmenter la durée de vie en bonne santé des patientes. Son efficacité thérapeutique a été démontrée à la fois sur la réduction des risques de développer une récurrence locorégionale et/ou une métastase à distance (29). Cette efficacité de la CTA (réduction relative du risque) serait globalement stable³⁰ entre les nombreux profils de tumeurs (31). D'après ces données agrégées, la CTA aurait une *efficacité* de 25 à 35 % sur la réduction des récurrences et de 10 à 20 % sur la réduction de la mortalité globale (31). En termes de *sécurité*, l'augmentation du risque absolu de décès par toxicité aiguë³¹ serait minime (0,2 - 0,3 %) mais variable selon le protocole reçu de CTA (31). Sur le peu d'études disponibles, le risque de complication grave et de décès par toxicité « différée » (leucémie, insuffisance cardiaque) serait également modéré (30).

► Implications pour démontrer l'utilité clinique d'une SG

Pour certaines patientes susceptibles de recevoir de la CTA mais à risque intermédiaire de récurrence (donc plutôt de bon pronostic), le bénéfice individuel de la CTA serait globalement modeste. L'existence d'incertitude résiduelle sur le bénéfice de la CTA, pour certains profils de cancer du sein et la préférence des professionnels de santé pour **des référentiels de prescription régionaux³² plus ou moins divergents entre eux**, font qu'il existe des situations susceptibles d'entraîner de l'hétérogénéité dans les décisions de CTA en contexte français (32, 33). Le recours à une SG permettrait d'optimiser cette décision en ne retenant cette option de traitement que chez les patientes dont le **bénéfice serait suffisant** (score génomique au-dessus du seuil de décision) et en l'omettant chez celles qui en subiraient principalement **les effets délétères** (score génomique au-dessous du seuil de décision) (2, 22, 25).

De nombreuses études visant à faire la preuve d'une valeur (ajoutée) pronostique ou prédictive des SG ont été minutieusement analysées dans le **rapport d'évaluation de l'INCa en 2013** pour les SG de première génération et dans d'autres rapports très récents à l'exemple de celui de l'INESS³³ d'octobre 2018 pour les SG de deuxième génération (34, 35). Ces institutions ont remonté la grande difficulté à pouvoir conclure avec confiance en faveur de l'utilité clinique d'une SG sur la base **d'études observationnelles** (de validation clinique). Partant d'échantillons de collections tumorales archivées couplées à certaines données cliniques, ces études explorent par de **multiples analyses rétrospectives** la présence de liens statistiques, d'interaction, de corrélations ou de différences pronostiques en vue d'apporter des preuves indirectes d'utilité clinique. En raison de nombreux risques de biais méthodologiques et de critères de jugement cliniquement moins pertinents, aucune balance bénéfice sécurité fiable et rassurante ne peut être établie en particulier dans le cas d'une maladie grave mais curable comme le cancer du sein. Les **études interventionnelles d'impact clinique³⁴** publiées ou en cours de réalisation servent à évaluer spécifiquement les nouvelles stratégies de prise en charge basées sur une SG. Si concluantes et appropriées, elles sont les seules susceptibles d'apporter des preuves directes d'utilité clinique.

³⁰ L'absence de recherche de HER2 et la faible représentation des tumeurs de grade 1 dans cette analyse pourraient nuancer la robustesse de ces constatations (30).

³¹ Estimée par la mortalité sans récurrence tumorale après un an de suivi dans les essais thérapeutiques (21)

³² Absence d'un document national commun concernant les indications et les non indications de CTA

³³ Institut d'excellence Nationale

³⁴ Essais prospectifs et comparatifs au mieux randomisés évaluant une nouvelle intervention

1.2 Caractéristiques techniques des SG en évaluation

Les tests Mammaprint et Oncotype Dx sont dits de première génération car ils ont été développés et validés exclusivement sur la génomique des tumeurs à partir de 2002 / 2004 pour guider la décision de CTA (36, 37). Les deux autres tests Prosigna et Endopredict, dits de deuxième génération, ont été développés et validés plus récemment en combinant la génomique des tumeurs à deux CCPT³⁵ à partir de 2009 / 2011 (38, 39). Un tableau synoptique des caractéristiques techniques des quatre SG est consultable en annexe 1.

1.3 Population ciblée préalablement pour l'évaluation

1.3.1 Critères clinico-pathologiques traditionnels (CCPT)

Avant de décider d'administrer ou non une CTA, la chirurgie d'exérèse doit être réalisée et carcinologiquement complète. En cas de cancer localisé du sein, les critères clinico-pathologiques traditionnels (CCPT) sont :

- Les critères pronostiques estimant le risque de récurrence tumorale :
 - l'âge de la patiente ;
 - le type histologique ;
 - la taille tumorale (T) ;
 - l'envahissement ganglionnaire (N) ;
 - le grade histopronostique ;
 - la présence d'embolies lymphovasculaires ;
 - le statut des récepteurs aux œstrogènes (RO) et à la progestérone (RP) ;
 - le statut des récepteurs HER2.
- Les critères prédictifs de l'efficacité d'une thérapie ciblée :
 - le statut immunohistochimique RO+ permet d'envisager un traitement par hormonothérapie adjuvante (HTA) ;
 - le statut immunohistochimique HER2 + permet d'envisager un traitement par une thérapie anti-HER2.

1.3.2 Populations ciblées : apport des recommandations les plus récentes

L'ensemble des grandes recommandations internationales préconise actuellement l'option d'utiliser d'une SG pour optimiser la décision de CTA (*cf.* Tableau 1). Concernant les critères d'éligibilité de la population ciblée, leur unique point de convergence est le statut immunohistochimique RO+/HER2- de la tumeur.

Le tableau récapitulatif ci-dessous (*cf.* Tableau 1) montre une grande hétérogénéité relative aux indications retenues et à la nature des tests à recommander. Ces divergences concernent la taille tumorale, le statut ganglionnaire ou de grade histopronostique. Ces documents sont imprécis sur les types histologiques à recommander. Il apparaît également que d'autres spécificités propres à certains pays (notamment en contexte nord-américain) puisse expliquer les différences notables observées par rapport au contexte français (40).

Pour toutes ces raisons, les recommandations internationales sont peu contributives en vue de l'élaboration du présent rapport.

³⁵ La taille tumorale (T) et le statut ganglionnaire (N)

Tableau 1 : Recommandations internationales et nationales relatives aux indications de SG

Documents	Formulation des recommandations
NCCN v2.1, 2018 (41)	Oncotype Dx ++ (plutôt à privilégier par rapport aux autres SG) Tumeurs RO+/HER2- pT1b à pT3, sans plus de précisions Si pN0 : recommandation forte Si patiente pN1mi ou N1a candidate à une CTA : recommandation faible
AJCC (TNM) 8th edition³⁶	Oncotype Dx ++ (plutôt à privilégier) Chez certaines patientes sélectionnées pT1-T2 RO+/HER2- pN0
Saint Gallen, 2017 (42)	Oncotype Dx, Prosigna, Mammaprint, Endopredict, BCI Indication privilégiée : incertitude décisionnelle RO+/HER2- pT1c-pT2 (T<3cm) pN0 de grade 2 et/ou Ki67 intermédiaire (sans consensus si pN1) Aucune indication chez les patientes non susceptibles de recevoir une CTA ou dont l'indication de CTA est évidente
European group on tumors markers (EGTM), 2017 (43)	Oncotype Dx, Prosigna, Endopredict : tumeurs RO+/HER2- pN0-1 Mammaprint : toute tumeur pN0-1 à risque clinique élevé (selon AOL modifié)
ASCO, 2016/2017 (22, 44)	Oncotype Dx, Prosigna, Endopredict : tumeurs RO+/HER2- pN0 (candidate à une CTA) Mammaprint : tumeurs RO+/HER2- pN0 et pN1 à « risque clinique élevé » (selon AOL modifié) (si information et accord de la patiente)
ESMO, 2015 (45)	Oncotype Dx, Mammaprint, Prosigna, Endopredict (si disponible) : tumeurs RO+/HER2- pN0-1 à risque intermédiaire Aucune indication en cas en dehors de cette catégorie de risque
Nice/Saint Paul de Vence, 2015 (46)	Aucune précision sur l'identité des SG à recommander Tumeurs RO+/HER2- pT1c-pT2 pN0 de grade 2

³⁶ Giuliano AE et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin. 2017 Jul 8;67(4):290-303.

1.3.3 Population ciblée par la HAS

► Population à risque intermédiaire susceptible de recevoir une CTA

Depuis 2009³⁷, les recommandations internationales concernant les critères de décision d'une CTA ont été largement impactées par l'introduction des SG (47-49).

Un certain nombre de recommandations, d'études pronostiques ou d'enquêtes de pratique indiquent que les tumeurs RO+/HER2- de taille infracentimétrique (pT1ab) et sans envahissement ganglionnaire (pN0) possèdent un excellent pronostic, expliquant que dans 93 à 98 % des cas, en France, ces patientes ne nécessitent³⁸ ni reçoivent de CTA (50-55). Chez ces patientes se pose avant tout un choix entre « absence de traitement systémique » et « HTA exclusive ». D'autres patientes encore possèdent une maladie localement avancée chez qui une CTA sera, sauf exception, systématique (56).

Il est donc possible de retenir une population préalablement ciblée à « risque intermédiaire » de récurrence pour qui un choix doit se faire au cas par cas entre une « HTA exclusive » et « HTA + CTA » (cf. Figure 1. ci-dessous).

Ces patientes sont celle possédant une tumeur RO+/HER2- (cf. Tableau 1 ci-dessus) pour qui une décision de CTA n'est ni systématique (patientes à risque élevé), ni trop peu fréquente (petite tumeur de faible risque ou de risque intermédiaire bas).

Cette population préalablement ciblée pour l'évaluation :

- doit avoir une tumeur de phénotype RO+/HER2- de stade pT1c-pT2 (taille tumorale entre 1 et 5 cm), opérable d'emblée³⁹ sans envahissement ganglionnaire (pN0) ou avec un envahissement limité (pN1mi/mic⁴⁰-pN1a) (cf. classification TNM en annexe 2)
- devrait être susceptible de recevoir une CTA.

La population ciblée susceptible de recevoir une CTA peut se scinder en deux sous populations distinctes sur la base des CCPT précédemment décrits :

- les patientes chez qui une indication de CTA était indiquée sur la base de critères de mauvais pronostic, où peut se poser la question d'une désescalade thérapeutique ;
- les patientes chez qui l'indication de CTA reste incertaine (situation d'incertitude décisionnelle), où se pose la question d'une optimisation de la décision finale.

Patientes chez qui existe une indication préalable de CTA

Historiquement, les essais thérapeutiques qui ont validé l'intérêt de la CTA dans le cancer du sein ont le plus souvent inclus des tumeurs avec envahissement ganglionnaire (tumeurs N+). Les essais de phase III ont progressivement validé l'intérêt thérapeutique d'une CTA pour les cancers pN0 de « mauvais pronostic ».

L'examen des critères d'inclusion des essais, des recommandations, des enquêtes de prescription en contexte français permettent de proposer cinq critères de mauvais pronostic majoritairement retenus pour décider d'une CTA chez les patientes à tumeur RO+/HER2- (51, 57, 58) :

- l'envahissement ganglionnaire (pN1) ;
- l'âge jeune de la patiente (< 35/40 ans) ;
- la taille tumorale pT2 (T >2-3 cm) ;

³⁷ Dès 2007/2008 avec les recommandations américaines de l'ASCO et du NCCN (40)

³⁸ Sauf certains cas particuliers plutôt rares de pT1b dits de plus « haut risque » de profil hétérogène et complexe. Aucune définition précise les concernant n'est retrouvée dans les recommandations

³⁹ C'est-à-dire sans indication de traitement néo-adjuvant en préopératoire (CTA ou HTA)

⁴⁰ Micrométastase ganglionnaire

- la présence d'emboles lymphovasculaires ;
- le grade 3 histopronostique.

Ces critères de mauvais pronostic sont fréquemment retenus dans les essais les plus récents de CTA⁴¹ à tumeur HER2-(59). Ils font tous partie des CCPT recommandés par la HAS dans son guide ALD (60). *A contrario*, bien que largement utilisé en contexte français en raison d'un financement non spécifique par l'Assurance maladie⁴², le Ki67 est encore d'interprétation empirique en l'absence de seuils consensuels pour les experts (58, 61-65). Ce test n'est pas retenu à l'heure actuelle comme étant un critère de décision pour une CTA (42). Des recommandations lui reconnaissent toutefois une valeur pronostique (42).

Patientes en situation d'incertitude décisionnelle

Cette sous-population d'intérêt en particulier semble conduire les professionnels à une hétérogénéité de pratiques (33). Des CCPT concordants mais équivoques ou un profil de discordance majeure entre ceux-ci peuvent générer ce genre de situation, d'autant plus fréquemment qu'augmente le nombre de CCPT à disposition des professionnels dans le cancer du sein (par exemple : Ki67, index mitotique...).

Dans ces situations délicates, la valeur ajoutée de chaque CCPT par rapport aux autres est recherchée dans un premier temps à l'aide de tests de corrélations et d'indépendances statistiques au sein d'un modèle multivarié explicatif (3, 66). Dans ces situations, le recours à un modèle multivarié de prédiction retenant l'ensemble des CCPT indépendants pour estimer le *risque de récurrence* d'une patiente donnée et son *bénéfice anticipé* vis-à-vis d'un traitement à l'efficacité connue est une solution clairement envisagée par certains professionnels (50, 67). Toutefois, il existe à ce jour peu de ces outils d'aide à la décision prenant en compte suffisamment de CCPT du cancer du sein. Deux modèles reviennent particulièrement dans la littérature : PREDICT⁴³ et Adjuvant ! Online⁴⁴. Leur utilisation a été récemment encouragée par l'*American Joint Committee on Cancer* (68). Toutefois, l'INCa ne préconise actuellement pas cette utilisation en routine pour décider d'une CTA (69). De nombreux travaux recommandent et encadrent le développement et la validation de ces modèles en vue d'être transférés dans la pratique (24, 70, 71).

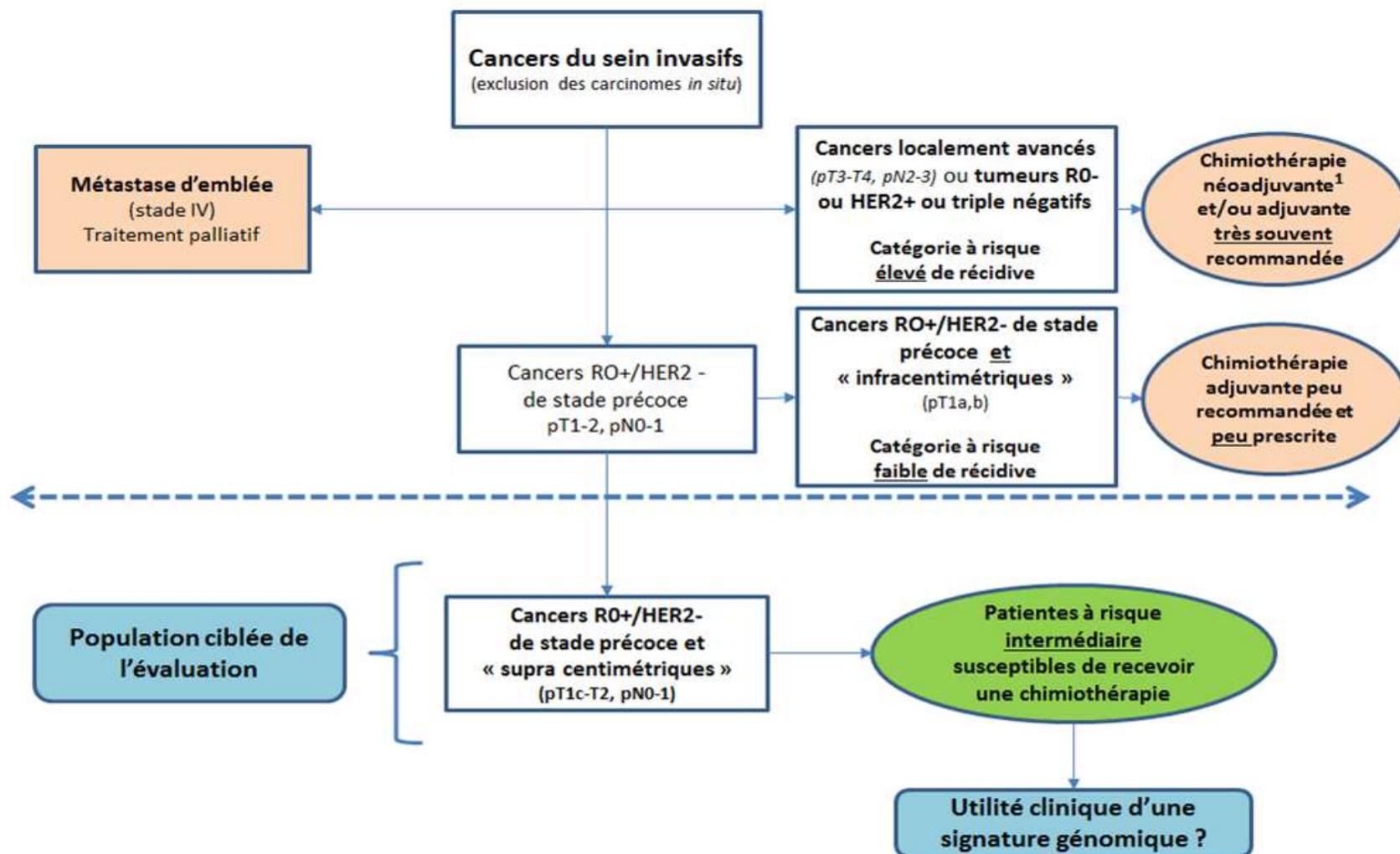
⁴¹ Par exemple: les essais thérapeutiques collaboratifs ABC : NSABP 49, NSABP 46, USOR 06090

⁴² Code CCAM : ZZQX045 : Examen immunocytochimique ou immunohistochimique de prélèvement cellulaire ou tissulaire fixé avec trois à cinq anticorps, avec quantification du signal pour chaque anticorps

⁴³ http://www.predict.nhs.uk/predict_v2.1/tool (site consulté le 02/10/2018)

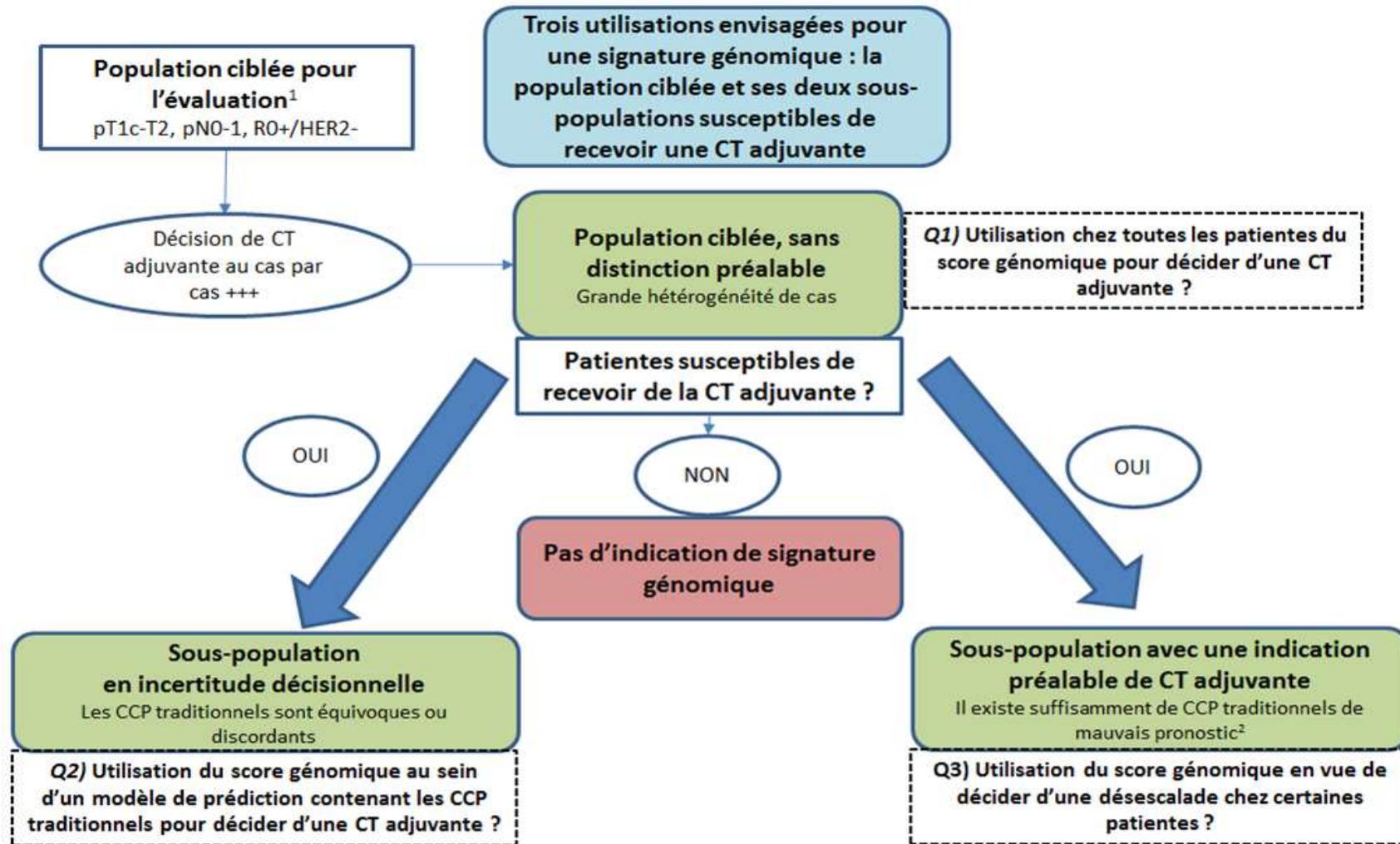
⁴⁴ Site non accessible le 02/10/2018

Figure 1 : Population ciblée préalablement pour l'évaluation susceptible de recevoir de la CTA



1- En complément d'une réduction du risque de récurrence, un traitement systémique néoadjuvant (chimiothérapie, hormonothérapie, antiHER2) adapté en fonction du profil tumoral (hormono-résistance, statut HER2) vise à améliorer le taux de conservation du sein et de résectabilité complète d'une tumeur localement avancée.

Figure 2. : Description des trois utilisations de SG dans la population préalablement ciblée susceptible de recevoir une CTA



1- Cancer de phénotype luminal (RO+) et HER2- de stade I-II (T < 5cm), opérable d'emblée sans indication de CT néoadjuvante

2- Critères tirés des référentiels existants, des études de pratique française et des critères d'inclusion dans les essais thérapeutiques de CT adjuvante en phase III

1.4 Évènements cliniques en contexte adjuvant

Dans le cadre d'un cancer localisé résecable à visée curative (contexte adjuvant), les événements cliniques pertinents au cours du suivi clinique sont (72, 73) :

- le décès : par évolution de la maladie cancéreuse⁴⁵, complications des traitements, ou toute autre cause retrouvée (accident, maladie chronique) ;
- la récurrence tumorale : locale (sein restant), régionale (paroi thoracique, ganglion) ou métastase à distance (extension à un autre organe) ;
- l'évènement indésirable grave (EIG) non létal lié au traitement ;
- le vécu de la patiente (symptômes, qualité de vie, satisfaction, anxiété/dépression).

1.5 Remontées d'utilisation en vie réelle (données RIHN 2016-2017)

La Direction générale de l'offre de soins (DGOS) a transmis à la HAS le volume d'environ 4 500 tests financés dans le cadre du Référentiel des actes Innovants Hors Nomenclature (RIHN) au cours de l'année 2017.

Depuis avril 2016, les quatre SG de cette évaluation sont financées de façon dérogatoire et transitoire par un acte global⁴⁶ du RIHN conditionné par « un recueil prospectif et comparatif de données cliniques ou médico-économiques »⁴⁷. Une étude prospective randomisée multicentrique de 800 patientes⁴⁸ avec pour objectif de comparer l'impact décisionnel des différentes SG sur la prescription de CTA avec un suivi clinique prévu de cinq ans (OPTIGEN – Promotion UNICANCER) était prévue. À ce jour, cette étude française multicentrique (60 centres) n'a toujours pas commencé.

À cette étude comparative s'est substitué plutôt un recueil indépendant par SG des caractéristiques des patientes testées et des données « d'impact décisionnel »⁴⁹ mais sans donnée de suivi clinique. Ces données descriptives sont déclaratives et ne prétendent pas à l'exhaustivité.

Les données du test Oncotype Dx ont été transmises par le fabricant (sur quatre mois de recueil) et par des acteurs académiques pour les tests Prosigna et Endopredict (16 à 19 mois de recueil). Concernant le test Mammaprint, le peu d'éléments transmis par le fabricant dans le cadre du financement RIHN était selon eux trop confidentiel (donc ils n'ont pas été rapportés dans ce rapport).

Les caractéristiques des patientes testées par une SG entre 2016 et 2017 dans le cadre du RIHN (condition d'utilisation en vie réelle) sont consultables en annexe 3.

Au total, peu d'éléments majeurs peuvent être déduits de ces données déclaratives, disparates et succinctes⁵⁰ (à l'exception du test Endopredict). Toutefois, il est à noter qu'une patiente déclarée sur cinq avait un envahissement ganglionnaire macrométastatique (pN+) pour les tests Oncotype Dx et Endopredict. L'utilisation d'une SG pour les tumeurs de grade 2 (cible potentiellement la plus intéressante) représentait respectivement 68 % (Oncotype Dx), 75 % (Endopredict) et 81 % (Prosigna) des prescriptions. Les types histologiques des tumeurs étaient hétérogènes (environ 20 % de tumeurs non canalaire). Le nombre des tests réalisés en dehors du référentiel REMAGUS de

⁴⁵ C'est-à-dire après une récurrence tumorale authentifiée

⁴⁶ Non soumis à accréditation mais réalisé au mieux dans un laboratoire de biologie accrédité selon la norme ISO 15189 (COFRAC) ou d'un laboratoire d'anatomopathologie impliqué dans une démarche d'évaluation externe de la qualité en vue d'une accréditation ISO 15189 (AFAQAP)

⁴⁷ <https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/recherche-et-innovation/rihn> (site consulté le 31 octobre 2018)

⁴⁸ 200 patientes pour chacune des quatre SG en évaluation

⁴⁹ Les données d'impact décisionnel transmises à la HAS n'ont pas été analysées car elles ne faisaient pas parties du champ d'évaluation de ce rapport (cf. critères de sélection des études du protocole)

⁵⁰ Aucune donnée n'a été transmise au-delà de juillet 2016 pour Oncotype Dx et elles se sont limitées à 40 % des tests réalisés d'une seule plateforme pour Prosigna (Institut Curie)

l'Institut Curie (74) a été significatif pour le test Prosigna. Cette particularité pourrait être expliquée par le caractère monocentrique (centre expert) des données recueillies pour Prosigna alors que les données ont été multicentriques et nationales pour le test Endopredict.

1.6 État des lieux des principales évaluations technologiques nationales

Les rapports d'évaluation technologiques des principales agences nationales d'évaluation et des groupes d'experts internationaux en langue française ou anglaise sur les cinq dernières années ont été consultés jusqu'en octobre 2018. Sont rapportés ci-dessous, les conclusions concernant l'existence ou non d'une démonstration de l'utilité clinique⁵¹ de certaines SG dans le cancer du sein :

- l'**INCa**⁵² (décembre 2013) ne peut conclure en faveur de l'utilité clinique⁵³ des tests Oncotype Dx et MammaPrint (34) ;
- l'**AHRQ** (Etats-Unis, mai 2014) ne peut conclure spécifiquement quant à l'utilité clinique⁵⁴ des tests Oncotype Dx et MammaPrint, faute de preuves directes apportées par des études d'impact clinique au moment de l'élaboration du rapport (75) ;
- le **CADTH** (Canada, 2014 et 2017) dans le cadre de deux évaluations rapides successives, ne peut conclure spécifiquement quant à l'utilité clinique des SG, notamment le test Oncotype Dx (76-78) ;
- l'**EGAPP** (Experts internationaux, décembre 2014) ne peut conclure quant à l'utilité clinique⁵⁵ des tests Oncotype Dx et MammaPrint, faute de preuves directes apportées par des études d'impact clinique et le très faible niveau de preuve *des autres études* au moment de l'élaboration du rapport (10) ;
- le **MSAC** (Australie, juillet et novembre 2017) n'autorise la prise en charge financière ni du test Oncotype Dx ni de Prosigna pour guider l'administration de CTA en raison d'une « valeur ajoutée incertaine » quant à leur utilité clinique par rapport aux CCPT, du comparateur proposé qui n'est pas approprié et peu standardisé. L'agence est dans l'attente des résultats d'études d'impact clinique (essai prospectif randomisé) comme TAILORx (79, 80). Pour les patientes en incertitude décisionnelle, l'agence recommande l'utilisation conjointe de modèle de prédiction du bénéfice de la CTA basés sur les CCPT (*AOL, PREDICT*), et d'une discussion multidisciplinaire prenant en compte le profil de la tumeur et de la patiente ;
- le réseau européen **EunethTA**⁵⁶ (janvier 2018) dont les conclusions ont été endossées récemment par le **KCE** (Belgique, février 2018) rapporte l'absence d'une démonstration de l'utilité clinique à cinq ans⁵⁷ avec le test MammaPrint par rapport à l'algorithme décisionnel AOL modifié (81, 82) ;
- l'**IQWiG** (Allemagne) conclut :
 - en **octobre 2016**, que l'étude d'impact clinique MINDACT ne permet pas de démontrer un bénéfice clinique avec la stratégie basée sur le test MammaPrint en raison d'un suivi trop court des patientes de cinq ans (83). Une publication plus récente des auteurs confirme selon eux l'absence de démonstration de l'utilité clinique de ce test (84) ;

⁵¹ Les analyses médico-économiques ont été exclues de cet état des lieux.

⁵² En collaboration avec la Société française de sénologie et de pathologies mammaires (SFSPM).

En l'absence d'une démonstration probante de « valeur ajoutée » pronostique ou prédictive des SG par rapport aux CCPT dans les études de validation

⁵⁴ Selon eux, la valeur ajoutée pronostique par rapport aux CCPT a toutefois été apportée pour les tests Oncotype Dx et MammaPrint mais avec un niveau de preuve « modéré » basé sur des méta-analyses d'association statistique (*hazard ratio*)

⁵⁵ Cinq revues systématiques ont été identifiées et analysées pour répondre à la question de l'utilité clinique des signatures.

⁵⁶ La HAS a été relecteur externe de la note de cadrage et du rapport d'évaluation

⁵⁷ Selon eux, l'analyse à dix ans est requise pour établir une conclusion définitive sur l'utilité clinique de ce test

- ▶ en **septembre 2018**, que l'étude TAILORx (d'après le résumé du rapport traduit en anglais⁵⁸) montre que le test Oncotype Dx peut permettre d'identifier certaines patientes en incertitude décisionnelle sans invasion ganglionnaire (pN0) chez qui il serait possible de se passer d'une CTA. Les patientes non susceptibles de recevoir une CTA, celles avec une indication préalable de CTA et celles pN+ ne sont pas éligibles au test, faute de donnée dans ces populations⁵⁹.
- le **NICE** (Royaume-Uni, décembre 2018) recommande dans son dernier rapport⁶⁰, comme une « option possible » l'utilisation des tests Endopredict, Oncotype Dx et Prosigna (mais pas Mammaprint) pour guider l'administration de la CTA chez les patientes RO+/HER2- pN0-N1mic, si :
 - ▶ il existe un risque intermédiaire de métastase estimé par un modèle de prédiction⁶¹ (PREDICT ou Nottingham Prognostic Index) ;
 - ▶ l'information apportée par la SG possède pour le clinicien une chance d'aider le choix thérapeutique en prenant en compte les préférences des patientes ;
 - ▶ les fabricants proposent leurs tests avec un allègement de prix ;
 - ▶ les cliniciens et les fabricants s'engagent à recueillir les données cliniques demandées par le NICE au sein d'un registre exhaustif et régulièrement mis à jour (National Cancer Registration).

1.7 Essais cliniques en cours

Les essais prospectifs et comparatifs en cours de réalisation susceptibles de démontrer l'impact et l'utilité cliniques d'une SG ont été colligés en annexe 4. Elles concernent les tests Oncotype Dx et Prosigna (85)

⁵⁸ <https://www.iqwig.de/en/press/press-releases/biomarker-tests-in-breast-cancer-new-study-data-indicate-advantage-for-certain-patients.10059.html> (consulté le 11 septembre 2018)

⁵⁹ L'IQWIG n'a pas réuni de GT multidisciplinaire pour définir ce qu'était un profil de population en incertitude décisionnelle. Il n'a pas non plus évoqué le manque de fiabilité des seuils de décision thérapeutique et la valeur ajoutée du test Il n'a pas discuté des caractéristiques de la population concrètement incluse dans TAILORx

⁶⁰ <https://www.nice.org.uk/guidance/dg34/chapter/1-Recommendations> (site consulté le 20 décembre 2018)

⁶¹ La HAS note que l'estimation du risque individuel de récurrence par un modèle de prédiction n'est actuellement ni recommandée par l'INCa et ni utilisée en contexte français. La HAS se demande également si une définition de ce qu'est « un risque intermédiaire de métastase » avec un modèle comme PREDICT est suffisamment bien établie dans la littérature.

2. Protocole de l'évaluation

2.1 Objectifs de l'évaluation

L'objectif principal de ce travail est d'évaluer l'utilité clinique de chacune des SG actuellement inscrites au RIHN dans leur revendication respective pour optimiser la décision de CTA.

L'objectif secondaire est de déterminer leur degré de concordance décisionnelle (homogénéité et transposabilité des décisions entre SG).

2.2 Méthode de l'évaluation

La méthode de cette évaluation (86) se fonde sur :

- 1) une analyse critique des données de la littérature après une recherche systématique conduite par la HAS ;
- 2) une consultation d'un groupe de travail multidisciplinaire constitué d'experts externes et d'une patiente en vue de recueillir leur position argumentée concernant la sélection et l'analyse de la littérature mais aussi certains éléments de compréhension relatifs au contexte français ;
- 3) un recueil du point de vue des parties prenantes (représentants d'organismes professionnels, association de patients) concernant la clarté et la lisibilité du rapport provisoire ainsi que leur appropriation de ce travail et les conséquences pratiques et organisationnelles qu'ils souhaitaient faire remonter ;
- 4) le recueil de la position institutionnelle de l'INCa.

2.3 Champ d'évaluation

2.3.1 Critères d'éligibilité

Le périmètre général de cette évaluation est limité aux SG et à la population d'intérêt répondant aux critères cliniques et médicotecniques ci-dessous (cf. Tableau 2).

Tableau 2 : Critères d'éligibilité cliniques et médicotecniques

Critères d'éligibilité cliniques	Critères d'éligibilité médicotecniques
Patiente en bon état général, sans comorbidités importantes.	Inscription au RIHN.
Cancer infiltrant et localisé du sein avec marges saines (R0).	Analyse de la pièce opératoire
Immunohistochimie : RO+/HER2- (\pm RP)	Score génomique interprétable selon un seuil de décision thérapeutique ⁶²
Stade précoce (pT1c-T2 ; pN0-N1, opérable d'emblée)	Les CCPT sont à disposition du clinicien

⁶² La capacité d'une SG à classifier une tumeur en sous type moléculaire sort du périmètre de cette évaluation.

Au total, la population ciblée préalablement pour l'évaluation se constitue de patientes en bon état général, sans comorbidités, atteintes d'un cancer du sein invasif opérable d'emblée, de phénotype R0+/HER2-, de stade pT1c-T2 (diamètre tumoral entre 1 et 5 cm), pN0-1 (envahissement ganglionnaire nul ou limité) qui devraient être susceptibles de recevoir une CTA sur la base des CCPT

2.3.2 Les trois utilisations pour guider une décision de CTA

Une SG peut être utilisée selon deux modalités différentes :

- une utilisation centrée exclusivement sur la valeur du score génomique avec **un objectif de stratification** des patientes en deux (ou trois) prises en charge distinctes selon un seuil de décision préétabli (seuil de décision thérapeutique), par exemple : risque génomique faible (abstention) et risque génomique élevé (administration du traitement) ;
- une utilisation avec **un objectif de prédiction du risque** selon un seuil de risque ou de bénéfice préétabli (70). Le score génomique est dans ce cas combiné aux CCPT dans un **modèle clinico-génomique** possédant un seuil de décision, par exemple : sous le seuil de décision thérapeutique (abstention), au-dessus du seuil de décision (administration du traitement).

En pratique, ces deux modalités d'utilisation se réfèrent concrètement à *trois situations cliniques distinctes* en contexte adjuvant :

- une situation où toutes les patientes de la population ciblée peuvent avoir une décision de CTA systématiquement et exclusivement basée sur leur risque génomique en vue de se substituer à la stratégie habituelle **(Q1)** ;
- une situation où seule les patientes en incertitude décisionnelle mais susceptibles de recevoir une CTA au sein de la population ciblée. Ces situations de décisions plus complexes devraient inclure l'ensemble des CCPT informatifs dans un modèle de prédiction multifactoriel performant dont l'intégration d'un score génomique⁶³ optimiserait encore la performance pronostique **(Q2)**. La décision finale devrait s'établir sur l'estimation du risque individuel et du bénéfice par le modèle complet (66, 70, 90, 91) ;
- une situation où les patientes avec une indication préalable de CTA sur la base des CCPT au sein de la population ciblée mais ayant un faible risque génomique pourraient au final ne pas en recevoir, le tout, sans préjudice clinique (désescalade) **(Q3)**.

2.4 Stratégie de recherche et de sélection des publications

La recherche systématique est structurée autour des éléments du Tableau 3 suivant :

Tableau 3 : Modalités de la recherche documentaire

Période de recherche	Du 01/01/2002 ⁶⁴ au 06/08/2018 (dernière veille bibliographique) Publications en langue anglaise et française
Bases de données consultées	<i>Medline, Cochrane Library, Center for Reviews and Dissemination databases</i> (équation de recherche en annexe 13)
Recherches complémentaires	Registre des essais cliniques ; bibliographie des publications ; publications transmises par les professionnels (experts/représentants d'organisme) ; identification libre à partir de moteurs de recherche internet.

⁶³ I) Démonstration d'une valeur ajoutée du score génomique au sein du modèle de base (87) ii) Evaluation de l'impact clinique du nouveau modèle clinico-génomique en comparaison du modèle de base (sans génomique tumorale) (8, 23, 88, 89).

⁶⁴ Année de la première publication importante concernant une SG dans le cancer du sein (36).

Les critères de sélection et d'analyse des données sont colligés dans le tableau suivant (cf. Tableau 4).

Tableau 4 : Critères de sélection des études

Critères de sélection des études	<p>Critères de sélection clinico- techniques :</p> <ul style="list-style-type: none"> • population préalablement ciblée pour l'évaluation (cf. 1.3.3) et en particulier ses deux sous-populations susceptibles de recevoir une CTA, celle en situation d'incertitude et celle avec indication préalable de CTA (cf. 1.3.3) ; • critères techniques de la SG comme définis dans le Tableau 2. <p>Critères de sélection méthodologiques :</p> <ul style="list-style-type: none"> • revues systématiques avec ou sans méta-analyses d'études d'impact clinique prospectives comparatives randomisées (à défaut, non randomisées) ; • études d'impact clinique prospectives, comparatives et randomisées (à défaut, non randomisées) avec un suivi minimal de cinq ans (à partir du diagnostic de cancer) ; • études prospectives et comparative de concordance décisionnelle. <p>Publications non retenues :</p> <ul style="list-style-type: none"> • revues systématiques et méta-analyses d'études de validation analytique et/ou clinique comparative ou non⁶⁵ ; • études de validation analytique et/ou clinique ; • études sur la classification en sous type moléculaire des tumeurs ; • études médico-économiques ; • recommandations de bonne pratique ; • éditoriaux et lettres à la rédaction ; • revues générales ; • résumés (<i>abstract</i>) ; • doublons.
---	---

La sélection documentaire est réalisée en trois phases successives : 1) sur titres et résumés, 2) lecture *in extenso* des publications éligibles, 3) sélection finale des publications retenues. Cette sélection documentaire est décrite dans un diagramme de flux (cf. Figure 3).

2.5 Critères d'évaluation de l'utilité clinique

Selon les standards méthodologiques internationaux, le rapport bénéfice/sécurité d'une nouvelle intervention nécessite de définir et de hiérarchiser les critères d'évaluation les uns par rapport aux autres selon leur pertinence clinique (92, 93). Les critères d'évaluation, comme la survie globale et le contrôle de la maladie cancéreuse (récidive tumorale), s'appuient sur les standards élaborés par des groupes d'experts dans le cancer du sein en contexte adjuvant : ICHOM, STEEP et DATECAN (73, 94, 95). Les critères d'évaluation du protocole devraient également être en conformité avec les exigences des agences de régulation du médicament⁶⁶, l'utilisation des SG avec pour objectif d'impacter les indications de CTA ayant reçu préalablement une autorisation de mise sur le marché (96, 97).

⁶⁵ Performances établies rétrospectivement partant d'échantillons issus d'anciennes collections tumorales couplées à l'existence de données cliniques recueillies prospectivement ou non

⁶⁶ La *Food and drug administration* (FDA) et l'agence européenne du médicament (EMA)

► Critères primaires du bénéfice

Le décès et la récurrence tumorale sont les principaux événements cliniques à prendre en compte pour quantifier le bénéfice d'une CTA.

Les critères d'évaluation du bénéfice de la stratégie basée une SG sont par ordre d'importance :

- la survie globale ;
- **la survie sans maladie**⁶⁷ (= critère de substitution accepté par les agences de régulation et les groupes d'experts).

► Critères primaires de sécurité

La sécurité clinique est évaluée par le taux de survenue des événements indésirables graves (EIG) aigus et/ou différés, mais n'entraînant pas la mort (grade 3-4). On entend par « différés », par exemple la toxicité cardiaque sous anthracycline ou les cancers secondaires chimio-induits. Les événements « non graves », c'est-à-dire ceux de grade 1 et 2 selon la dernière grille du NCI-CTCAE, ne sont pas considérés comme suffisamment pertinents au cours d'un essai thérapeutique (cf. annexe 5).

► Balance bénéfice/sécurité : bénéfice net

L'objectif clinique d'une stratégie décisionnelle basée sur la génomique tumorale est d'atteindre une meilleure balance bénéfice/sécurité que sans ce nouveau marqueur. Le bénéfice net de la stratégie basée sur la SG doit être positif, statistiquement significatif, fiable méthodologiquement et cliniquement pertinent.

- 1) L'utilité clinique d'une SG réside au mieux dans une amélioration de la survie globale.
- 2) Si cette démonstration n'est pas observée (immaturité du suivi ou risque de récurrence trop faible), la balance bénéfice sécurité peut être établie sur un critère de substitution tel que la survie sans maladie.

Le bénéfice net d'une stratégie décisionnelle est établi entre :

- la différence absolue (Δ delta) de survie sans maladie ;
- la différence absolue (Δ delta) de toxicité grave imputable à une CTA.

► Critères secondaires d'évaluation

Contrôle du cancer :

Pour peu que l'approche statistique prenne en compte les risques compétitifs⁶⁸ de la patiente, la survie sans métastase à distance⁶⁹, le délai avant une première récurrence ou le délai avant une première métastase à distance⁷⁰ sont des critères pronostiques importants d'incurabilité et de mortalité (102-104). Néanmoins, le délai d'apparition d'une métastase en contexte adjuvant (particulièrement en cas de stade précoce) est peu fréquent. La prise en compte exclusive des métastases, sans les autres événements cliniques délétères et pertinents (récurrences locorégionales, autres cancers) pour estimer la balance bénéfice sécurité d'une CTA, n'est jamais observée dans les études. À cet effet, ces critères ne sont pas recommandés en tant que critères primaires pour

⁶⁷ Critère clinique composite et pragmatique, marqueur du contrôle de la maladie cancéreuse, il est largement recommandé pour estimer l'impact d'un nouveau traitement en contexte adjuvant (94-100). Il intègre comme premier événement le décès avant récurrence, toute récurrence tumorale, un second cancer invasif métachrone (sein controlatéral ou sur autre localisation). L'intégration d'un second cancer invasif se justifie par le fait que qu'il représente un événement clinique grave capable d'impacter fortement le pronostic vital du patient ou être une complication potentielle de la CTA (cancer chimio-induit) ou par ailleurs être difficilement distinguable de métastase à distance (ce qui limite le risque de sous-estimer le nombre de métastases à distance)

⁶⁸ Par exemple par l'estimation de l'incidence cumulée spécifique des métastases à distance (101)

⁶⁹ Ce critère composite n'intègre pas la première récurrence locorégionale et le second cancer invasif

⁷⁰ Ce critère n'intègre pas le décès avant récurrence et le second cancer invasif

évaluer les résultats d'un essai thérapeutique en contexte adjuvant (94, 96-98, 100). Ces critères intermédiaires exposent également à des risques sérieux de biais⁷¹ lors de l'analyse statistique (94, 95).

Qualité de vie :

Toute récurrence tumorale entraîne un impact délétère sur l'état psychologique et la qualité de vie de la patiente (105). Dans le contexte d'une maladie grave mais curable, l'évolution de la qualité de vie et du bien-être des patientes durant la phase de traitement restent des critères cliniques secondaires. De plus, ils sont plus sensibles au risque de biais que le contrôle de la maladie (critère primaire) si l'étude est en ouvert⁷² (93). La qualité de vie peut être mesurée à l'aide d'échelles internationales génériques (EQ-5D, SF-36) ou spécifiques du cancer (EORTC-QLQ-C30, FACT-G) ou bien spécifique du cancer du sein (QLQ-BR23). La qualité de vie peut être fluctuante au cours du temps, le moment le plus pertinent à son analyse peut être complexe à déterminer et le seuil cliniquement pertinent signifiant une altération de la qualité de vie peut être difficile à considérer (106, 107). En outre, la qualité de vie devrait être recueillie par un investigateur ou par un patient *en insu* du traitement reçu pour rendre ce critère subjectif informatif (108). Enfin, une altération de la qualité de vie au cours du suivi peut être liée à d'autres causes que la CTA : la chirurgie ou la radiothérapie.

Pour toutes ces limites d'interprétation, la détérioration de la qualité de vie est considérée comme un critère d'évaluation secondaire en contexte adjuvant.

Critères non cliniques :

D'autres indices non cliniques sont importants. Il y a le degré de concordance décisionnelle entre différentes SG pour un même échantillon de patientes et également la proportion d'adhésion (ou de compliance) des individus concernés par la décision thérapeutique du test en évaluation (ici, la SG).

2.5.1 Questions cliniques de l'évaluation (Q1 à Q4)

Pour chacune des trois situations cliniques identifiées, une question clinique se pose (Q1 à Q3). Chaque question est accompagnée d'un tableau synoptique (au format PICOTS⁷³) structurant son cadre d'analyse. La dernière question clinique évalue l'interchangeabilité des SG entre elles basée sur la transposabilité des décisions (Q4).

► Question d'évaluation Q1

Q1 : Dans la population préalablement ciblée pour l'évaluation (sans distinction), l'utilisation systématique d'une SG avec un seuil de décision thérapeutique préétabli pour administrer une CTA par rapport aux CCPT améliore-t-elle suffisamment le devenir clinique des patientes pour pouvoir se substituer à la stratégie habituelle ?

⁷¹ Certains événements compétitifs importants peuvent être censurés ou ignorés par le statisticien

⁷² Néanmoins, la qualité de vie est d'importance majeure en contexte métastatique (80)

⁷³ *Population, Intervention, Comparators, Outcomes, Time frame, Study design.*

Tableau 5 : Tableau PICOTS de la question d'évaluation Q1

Population cible	Toutes les patientes de la population préalablement ciblée (sans distinction) recevant une SG comme définie dans le Tableau 2.
Intervention à tester	Utilisation de la SG selon une stratégie décisionnelle guidée par la valeur du score génomique à interpréter selon son seuil de décision thérapeutique. Une valeur au-dessus du seuil de décision (haut risque génomique) oriente vers une CTA + HTA et une valeur au-dessous du seuil de décision (bas risque génomique) oriente vers une HTA exclusive.
Comparateurs existants	Utilisation des CCPT pour décider d'une CTA selon la stratégie habituelle : <ul style="list-style-type: none"> décision basée sur des recommandations de bonne pratique (critères de mauvais pronostic) ; décision basée sur un modèle de prédiction du risque n'utilisant aucun score génomique (par exemple : AOL ou PREDICT).
Critères d'évaluation	<p>Critères primaires (bénéfice/sécurité) :</p> <ul style="list-style-type: none"> survie globale ; survie sans maladie ou sans récurrence ; EIG grade 3 ou 4 lié à la toxicité de la CTA (selon NCI-CTCAE v5.0). <p>Critères secondaires :</p> <ul style="list-style-type: none"> survie sans métastase à distance ; délai avant une première récurrence ; délai avant une première métastase à distance ; altération du score de qualité de vie : échelles multidimensionnelles génériques EQ-5D, SF-36) spécifiques du cancer (EORTC-QLQ-C30, FACT-G) ou spécifique du cancer du sein (QLQ-BR23).

► **Question d'évaluation Q2**

Q2 : Chez les patientes susceptibles de recevoir une CTA en raison d'une incertitude décisionnelle au sein de la population préalablement ciblée pour l'évaluation, l'ajout d'un score génomique au sein d'un modèle de prédiction incluant les CCPT (algorithme décisionnel) pour la décision de CTA améliore-t-il suffisamment le devenir des patientes par rapport au même modèle de prédiction mais sans recours à un score génomique ?

Tableau 6 : Tableau PICOTS de la question d'évaluation Q2

Population cible	Sous-groupe en situation <u>d'incertitude décisionnelle</u> de la population préalablement ciblée susceptible de recevoir de la CTA recevant une SG comme définie dans Tableau 2.
Intervention à tester	Utilisation d'une SG en intégrant sa valeur génomique dans un modèle de prédiction du risque de récurrence ⁷⁴ avec un seuil de décision thérapeutique spécifique au modèle préétabli pour décider d'une administration de CTA.
Comparateurs existants	Utilisation des CCPT dans le même modèle de prédiction du risque mais n'incluant pas de score génomique

⁷⁴ Combinant le score génomique avec les CCPT

Population cible	Sous-groupe en situation <u>d'incertitude décisionnelle</u> de la population préalablement ciblée susceptible de recevoir de la CTA recevant une SG comme définie dans Tableau 2.
Critères d'évaluation	<p>Critères primaires (bénéfice/sécurité) :</p> <ul style="list-style-type: none"> • survie globale ; • survie sans maladie ou sans récurrence ; • EIG grade 3 ou 4 lié à la toxicité de la CTA (NCI-CTCAE v5.0). <p>Critères secondaires :</p> <ul style="list-style-type: none"> • survie sans métastase à distance ; • délai avant une première récurrence ; • délai avant une première métastase à distance ; • altération du score de qualité de vie : échelles multidimensionnelles génériques (EQ-5D, SF-36), spécifiques du cancer (EORTC-QLQ-C30, FACT-G), spécifique du cancer du sein (QLQ-BR23).

► **Question d'évaluation Q3**

Q3 : Chez les patientes avec une indication établie de CTA basée sur les CCPT au sein de la population préalablement ciblée pour l'évaluation, l'utilisation d'une SG pour procéder à une désescalade de CTA (en cas de valeur en dessous du seuil de décision thérapeutique) permet-elle d'améliorer le devenir clinique des patientes ?

Tableau 7 : Tableau PICOTS de la question d'évaluation Q3

Population cible	Sous-groupe <u>avec une indication établie de CTA</u> de la population préalablement ciblée susceptible d'en recevoir et disposant d'une SG comme définie dans Tableau 2.
Intervention à tester	Utilisation d'une SG de la manière suivante : désescalade de CTA si score génomique en-dessous du seuil de décision thérapeutique préétabli.
Comparateurs existants	Utilisation des CCPT qui retrouve des critères de mauvais pronostic motivant une décision préalable de CTA dans le cadre de recommandations ou d'un modèle de prédiction du risque de récurrence.
Critères d'évaluation	<p>Critères primaires (bénéfice sécurité) :</p> <ul style="list-style-type: none"> • survie globale ; • survie sans maladie ou sans récurrence ; • EIG grade 3 ou 4 lié à la toxicité de la CTA (NCI-CTCAE v5.0). <p>Critères secondaires :</p> <ul style="list-style-type: none"> • survie sans métastase à distance • délai avant une première récurrence ; • délai avant une première métastase à distance ; • altération du score de qualité de vie : échelles multidimensionnelles génériques (EQ-5D, SF-36), spécifiques du cancer (EORTC-QLQ-C30, FACT-G par exemple), spécifique du cancer du sein (QLQ-BR23).

► Question d'évaluation Q4

Q4 : Les décisions de CTA sont-elles concordantes entre les différentes SG ?

Les études prospectives et comparatives comparant entre elles plusieurs SG sur le même échantillon seront retenues. Les analyses rétrospectives sur des d'anciens échantillons tumoraux archivés ne sont pas acceptées pour leur nombreuses limites méthodologiques.

Seront extraits :

- l'indice de concordance décisionnelle (score Kappa) ;
- la proportion de changements décisionnels (en %).

2.6 Groupe de travail multidisciplinaire

2.6.1 Constitution

Les spécialités suivantes ont été sollicitées pour participer à cette évaluation :

- chirurgie oncologique du sein ;
- oncologie médicale ;
- biologie médicale et moléculaire ;
- anatomopathologie ;
- méthodologie et biostatistiques appliquées à la médecine (spécialité : oncologie).

Un appel à candidature pour les professionnels de ces spécialités et les patientes concernées a également été diffusé sur le site internet de la HAS.

Le groupe de travail (GT) s'est constitué de professionnels de santé proposés par les organismes professionnels suivants :

- Conseil national professionnel (CNP) de pathologie ;
- CNP de biologie médicale ;
- CNP d'oncologie médicale ;
- CNP de génétique clinique, chromosomique et moléculaire ;
- Société française de chirurgie oncologique (SFCO) ;
- Société française de médecine prédictive et personnalisée (SFMP) ;
- Société française de sénologie et pathologie mammaire (SFSPM) ;
- Groupe francophone de cytogénomique oncologique (GFCO).

Il a été complété par un appel à candidature sur le site internet de la HAS et par la candidature d'une patiente appartenant à l'association Vivre comme Avant.

Le CNP de gynécologie-obstétrique et l'Association des épidémiologistes de langue française (ADELF) ont également été sollicités mais n'ont proposé aucun nom d'experts.

En l'absence de proposition, la HAS a donc dû solliciter directement un certain nombre d'experts français spécialisés en méthodologie et biostatistiques appliquées à l'oncologie exerçant dans des unités de recherche sur la base de leurs titres et travaux.

Les associations de patients suivantes : « Ligue nationale contre le cancer », « Essentielles » et « Collectif K - Atout cancer » ont été enfin sollicitées mais n'ont pas proposé de noms.

2.6.2 Composition

Dans l'optique d'avoir toutes les disciplines équitablement représentés, les experts retenus pour les réunions du GT ont été :

- M. le Dr Laurent ARNOULD, Anatomopathologie, CLCC de Bourgogne, DIJON (21) ;

- Madame le Docteur Carine BELLERA, Méthodologie et biostatistiques⁷⁵, Institut Bergonié, BORDEAUX (33) ;
- Monsieur le Docteur Jean Marie BOHER, Méthodologie et biostatistiques, Institut Paoli-Calmettes, MARSEILLE (13) ;
- Madame le Docteur Patricia DE CREMOUX, Biologie médicale, CHU Saint Louis, PARIS (75) ;
- Monsieur le Docteur Frédéric FITENI, oncologie médicale, CHU de Nîmes, NÎMES (30) ;
- Madame le Docteur Sophie GOURGOU, Méthodologie et biostatistiques ; Institut du cancer de Montpellier, MONTPELLIER (34)
- Madame le Docteur Séverine GUIU, Oncologie médicale, Institut du cancer de Montpellier, MONTPELLIER (34) ;
- Monsieur le Docteur Jonathan LOPEZ, Biologie médicale, CHU LYON SUD, LYON (69) ;
- Madame le Docteur Elisabeth LUPORSI, Oncologie médicale, CHR Metz-Thionville, METZ (57) ;
- Monsieur le Docteur Gaëtan MAC GROGAN, Anatomopathologie, Institut Bergonié, BORDEAUX (33) ;
- Monsieur le Dr Jean-Baptiste SOURAUD, Anatomopathologie, Hôpital des Armées de Begin, SAINT MANDE (94) ;
- Madame le Docteur Christine TUNON DE LARA, Chirurgie oncologique du sein, Institut Bergonié, BORDEAUX (33) ;
- Madame Véronique WOLF, patiente, Association Vivre comme avant, LILLE (59).

Mme le Dr Alexia SAVIGNONI, Méthodologie et biostatistiques, Institut Curie, PARIS (75) a été excusée pour raison personnelle. M. le Dr Gaëtan MAC GROGAN a été excusé lors de la seconde réunion pour raison personnelle.

2.6.3 Déclaration d'intérêts

Les déclarations publiques d'intérêts (DPI) des membres du GT ont été analysées par la HAS en comité de validation des DPI, instance présidée par un déontologue (D. LUDET). Ces décisions sont toutes en conformité avec le guide des déclarations d'intérêts et de gestion des conflits d'intérêts⁷⁶ de la HAS. Selon le comité de validation, aucun des membres du GT n'a déclaré de liens d'intérêts majeurs en relation avec le sujet de cette évaluation. Les DPI sont consultables sur le site unique : DPI.sante.gouv.fr.

2.6.4 Recueil de la position argumentée du groupe de travail

Le groupe de travail s'est réuni le **02 mai 2018** pour émettre une position argumentée d'experts sur la version provisoire du protocole d'évaluation rédigé par la HAS en proposant des modifications et en répondant à des questions médicales sur la pratique française.

Le groupe de travail s'est de nouveau réuni le **02 juillet 2018** pour émettre une position argumentée et des commentaires sur l'analyse de la littérature retenue par la HAS concernant l'utilité clinique des SG. Les comptes rendus consolidés des réunions ont été validés par l'ensemble des membres du GT et sont présentés en annexe 10 et annexe 11. Les experts ont par ailleurs accepté que leur nom figure dans ce rapport.

⁷⁵ INSERM CIC-EC 14.01 (Epidémiologie Clinique), 146 rue Léo Saignat 33000 Bordeaux, France et INSERM, ISPED, Centre INSERM U1219 Bordeaux Population Health Center, Equipe Epicene, 146 rue Léo Saignat 33000 Bordeaux, France

⁷⁶ Haute Autorité de Santé. Guide des déclarations d'intérêts et de gestion des conflits d'intérêts Saint-Denis La Plaine : HAS, 2013. Disponible sur http://www.has-sante.fr/portail/upload/docs/application/pdf/guide_dpi.pdf.

La partie 4 du guide a été modifiée par décision du collège le 15 mars 2017

2.7 Recueil du point de vue des parties prenantes

Certaines associations professionnelles directement concernées par nos travaux et deux associations de patients ont été sollicitées en tant que partie prenante au sens du décret n°2013-413 du 21 mai 2013⁷⁷ par un courrier adressé à chaque président accompagné d'une version provisoire du rapport d'évaluation et d'un questionnaire. Ces retours devaient être le reflet du point de vue collectif de l'organisme ou de l'association dans sa subjectivité. Ces éléments d'appréciation n'avaient pas vocation à rechercher un consensus ni à remettre en question la qualité scientifique et l'indépendance des travaux de la HAS mais visaient à améliorer la lisibilité et la clarté du document final. Cette consultation visait également à éclairer le décideur sur le degré d'appropriation de ce travail par le terrain et sur les conséquences pratiques et organisationnelles que souhaitaient faire remonter ces acteurs.

Les associations professionnelles sollicitées ont été :

- CNP d'anatomopathologie ;
- CNP de biologie médicale ;
- CNP d'oncologie médicale ;
- Société française de sénologie et pathologie mammaire (SFSPM) ;
- Société française de chirurgie oncologique (SFCO) ;
- Société française de médecine prédictive et personnalisée (SFMPP) ;
- Groupe francophone de cytogénomique oncologique (GFCO).

Les deux associations de patients sollicitées et disposant d'un agrément sont :

- la Ligue nationale contre le cancer (LNCC) ;
- l'association « Vivre comme avant ».

2.8 Consultation de l'Institut national du cancer

Dans le même temps, le rapport provisoire a également été adressé à l'Institut National du Cancer (INCa) en relecture afin de recueillir ses commentaires et sa position institutionnelle comme organisme public référent dans le domaine du cancer.

⁷⁷ JO RF n°0116 du 22 mai 2013 page 8405, texte n°5.

3. Résultats de l'évaluation

3.1 Diagramme de flux de la recherche systématique

La Figure 3 ci-dessous résume la sélection bibliographique mise en œuvre dans ce rapport :

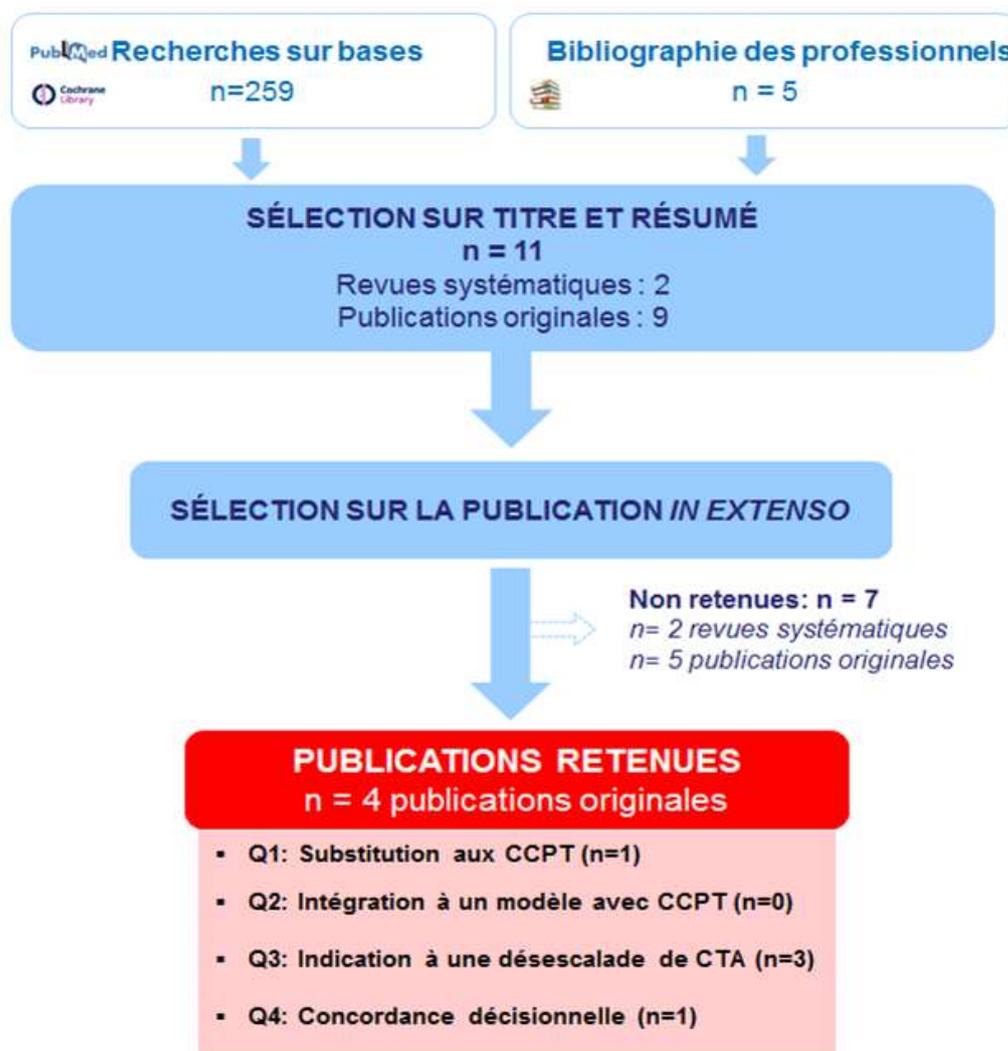


Figure 3 : diagramme de flux de la sélection bibliographique

Les références des quatre publications retenues pour le rapport sont :

- les résultats à cinq ans de l'essai MINDACT (109) ;
- les résultats à cinq et neuf ans de l'essai TAILORx (110) ;
- les résultats à cinq ans d'analyses ancillaires en marge de l'essai PLAN B (111) ;
- les résultats de concordance décisionnelle en phase de faisabilité préliminaire de l'essai OPTIMA (112).

Après lecture *in extenso*, les cinq publications originales non retenues sont :

- l'étude RASTER qui est une cohorte observationnelle monobras d'impact décisionnel de type « avant/après » pour qui la décision finale de CTA ne dépendait qu'en partie du risque génomique (113, 114) ;
- la publication partielle de TAILORx pour les patientes à risque génomique faible (115) ;

- une publication doublon des résultats de l'étude OPTIMA Prelim (116) ;
- la publication des résultats préliminaires à trois ans de l'analyse ancillaire en lien avec l'essai PLAN B (117).

3.2 Analyse des données de la littérature

3.2.1 Revues systématiques

La recherche de revues systématiques (RS) n'a pas permis d'identifier de méta-analyse évaluant l'impact clinique des SG. Aucune autre RS identifiée n'a été retenue dans la présente évaluation pour les raisons suivantes :

- RS d'études observationnelles (pro-)rétrospectives de validation clinique des performances pronostiques et/ou prédictives ;
- RS d'études d'impact décisionnel avant-après ;
- RS d'études médico-économiques ;
- RS demandant une mise à jour (absence des résultats finaux de TAILORx) ;
- RS doublon d'un rapport d'évaluation technologique.

3.2.2 Question 1 : Démontrer l'utilité clinique d'une décision de CTA basée sur une stratégie génomique par rapport à la stratégie habituelle en vue d'une substitution des CCPT

► Objectif médical

Décider d'une CTA chez les patientes de la population ciblée préalablement pour l'évaluation en se basant systématiquement sur le seuil d'intervention d'une SG prise isolément (si haut risque génomique → CTA ; si bas risque → abstention de CTA/HTA exclusive).

► Sélection bibliographique

La recherche systématique a permis de retenir une seule étude d'impact clinique comparative évaluant Mammaprint (109) qui prévoyait une double randomisation des traitements alloués (HTA exclusive ou association d'une CTA/HTA) entre les deux sous-groupes de patientes ayant un risque discordant entre la génomique et les CCPT (MINDACT, NCT00433589).

Aucune étude prospective et comparative d'impact clinique répondant à la question 1 n'a été identifiée avec Oncotype Dx, Prosigna et Endopredict.

► Principaux résultats de MINDACT (cf. annexe 6)

La population globale de l'étude (n= 6 693) était composée d'une population plus large et moins bien définie que celle ciblée pour l'évaluation c'est-à-dire que tous les profils immunohistochimiques confondus étaient représentés. Environ 19 % des tumeurs ne correspondaient à aucune indication actuelle de SG⁷⁸ et à un phénotype plus agressif. Ces patientes présentaient une indication de CTA dans 50 % des cas (AOL modifié indiquant un haut risque clinique).

Dans l'étude MINDACT, la démonstration d'une optimisation décisionnelle avec la stratégie basée sur la génomique (Mammaprint) devait s'appuyer sur l'analyse exclusive des patientes à risque discordant entre la génomique et la clinique (AOL modifié) correspondant à environ 33 % de l'effectif testé⁷⁹ (patientes C Bas/G Haut ou C Haut/G Bas). Toutes ces patientes étaient randomisées pour décider de l'allocation d'une HTA exclusive ou d'une CTA/HTA.

⁷⁸ Tumeurs de phénotype non hormonodépendant (RO+), triple négatifs ou HER2 + (Cf. Contexte)

⁷⁹ n= 2 187 patientes

Peu nécessaires au présent raisonnement, les principales limites de l'étude ne sont pas abordées dans cette partie mais un peu plus loin pour répondre à la question Q3. En effet, les résultats attendus d'impact clinique permettant de valider dans la population ciblée pour l'évaluation (tumeurs RO+/HER2- pN0 ou pN1) une supériorité de la stratégie génomique (avec Mammaprint) sur la stratégie clinique n'ont pas été fournis par les auteurs.

A la place, il est fourni une analyse comparant l'impact clinique de la stratégie génomique à celle clinique (en annexe 6) dans la population totale sur un critère de jugement inadéquat en contexte adjuvant (survie sans métastase à distance- SSMD). Les résultats ne montrent pas d'amélioration du pronostic des patientes avec Mammaprint, et même un gain minime avec la stratégie clinique (+ 0,3 % de SSMD à cinq ans). Les autres résultats de comparaison proposés par les auteurs dans la population ciblée de l'évaluation dans les suppléments de la publication (cf. annexe 6) ne sont pas pertinents au regard de la question posée⁸⁰.

► Conclusion

L'absence de résultats pertinents proposés par les auteurs dans la population ciblée pour l'évaluation ne permet pas de démontrer que la stratégie génomique avec le test Mammaprint conduit à une amélioration du devenir clinique des patientes par rapport à la stratégie clinique en vue d'une substitution.

3.2.3 Question 2 : Démontrer l'utilité clinique d'une décision de CTA basée sur un modèle de prédiction incluant un score génomique en association avec les CCPT dans la population ciblée en situation d'incertitude ?

► Objectif

Décider d'une CTA chez les patientes de la population ciblée en incertitude décisionnelle (par exemple : critères équivoques, discordance clinico-histo-pronostique) en se basant sur le seuil d'intervention thérapeutique d'un modèle de prédiction du risque et du bénéfice d'une CTA (risque estimé ou bénéfice anticipé au-dessus du seuil de décision → CTA).

► Sélection bibliographique

La sélection n'a pas permis d'identifier de publication susceptible de démontrer l'utilité clinique d'un score génomique si celui-ci est intégré à un modèle de prédiction de référence recourant aux CCPT.

► Conclusion

La démonstration de l'utilité clinique de combiner de la génomique tumorale aux CCPT au sein d'un modèle de prédiction ne peut être établie au vu de l'absence d'études susceptibles d'apporter une preuve directe de ce type, en particulier dans la population ciblée en incertitude décisionnelle.

3.2.4 Question n°3 : démontrer l'utilité clinique d'une SG pour décider d'une désescalade de CTA dans la population ciblée avec une indication préalable sur la base des CCPT

► Objectif

Décider d'une désescalade de CTA en toute sécurité en cas de risque génomique faible avec une SG chez des patientes de la population ciblée ayant une indication préalable de CTA sur la base des CCPT.

⁸⁰ Les auteurs ont fourni des données « après correction des erreurs de risque » en tableau S7 et S8 de la publication. Ils ont fusionné le pronostic des patientes sous HTA exclusive avec celles sous CTA/HTA (tableau S8 de la publication) ou ont étudié le pronostic des patientes sous la stratégie génomique sans la comparer à la stratégie clinique dans le tableau S9 de la publication.

► Sélection bibliographique

La recherche systématique a permis de retenir les résultats finaux de trois études d'impact clinique comparatives dont l'objectif était de valider l'option de désescalade de CTA :

- l'étude MINDACT (NCT00433589) avec une randomisation des traitements alloués entre HTA exclusive et CTA/HTA évaluant le test Mammaprint dans le sous-groupe des patientes à risque clinique élevé (C-haut) en discordance avec un bas risque génomique (G-bas) (109) ;
- l'étude TAILORx (NCT00310180) avec une randomisation des traitements alloués HTA exclusive ou CTA/HTA chez les patientes du sous-groupe à risque génomique intermédiaire (score RS entre 11-25) pour le test Oncotype Dx (110) ;
- des données prospectives ancillaires à l'essai PLAN B (NCT01049425) (111).

Aucune étude prospective et comparative répondant à la question 3 n'a été identifiée avec les tests Prosigna et Endopredict

► Principaux résultats des études retenues

MINDACT : Mammaprint (résultats cf. annexe 6)

L'objectif primaire de l'étude MINDACT était de démontrer la capacité de Mammaprint à permettre une désescalade de CTA en toute sécurité chez des patientes à haut risque de récurrence (C-haut) sur la base de l'algorithme AOL modifié. Dans ce sous-groupe de patientes à risque discordant (C-haut/G-bas) représentant 23 % de l'effectif de l'étude, toutes les patientes avaient selon les auteurs une indication préalable de CTA. Cette population était également plus large et plus hétérogène que la population ciblée par la HAS pour les mêmes raisons que précédemment (Q1).

La démonstration d'un pronostic qui restait favorable à cinq ans chez ces patientes laissées sous HTA exclusive était basée, selon les auteurs, sur un pronostic supérieur au seuil théorique de 92 % de survie sans métastase à distance (SSMD). L'objectif primaire de l'étude a donc été atteint avec une SSMD de 94,7 %, intervalle de confiance (IC) à 95 % [92,5 – 96,2], abstraction faite des nombreuses limites méthodologiques constatées à la lecture de l'étude⁸¹.

Il convient cependant de souligner :

- la comparaison était faite indirectement contre une valeur théorique (suivi longitudinale non comparatif des patientes). Elle est faite par rapport à un pronostic à cinq ans de 92 % sur la fréquence d'apparition d'une métastase fixée par les auteurs. Ce seuil semble permissif et n'est pas justifié par les auteurs. Il ne permet pas de décréter un bon pronostic à ces patientes laissées sous HTA exclusive. À titre de comparaison, les patientes de bon pronostic de TAILORx ne présentaient aucune métastase à distance dans 98 à 99 % des cas à cinq ans ;
- la démonstration d'une efficacité conservée de la CTA en intention de traiter⁸² et en per protocole (HR≈ 0,60 - 0,70) avec un bénéfice cliniquement pertinent (≥ 3 %) et statistiquement significatif sur la survie sans maladie (SSM) à cinq ans, critère d'évaluation recommandé en contexte adjuvant (84). Pour les autres critères de jugement (pour lesquels les événements cliniques sont plus rares), les tailles d'effet restent toutes du même ordre (bien que statistiquement non significatives) ;
- la valeur prédictive d'efficacité de la CTA concernant le test Mammaprint n'est pas démontrée au vu des données publiées (estimation permise par la double randomisation des tests discordants (14)). L'efficacité de la CTA paraît même supérieure chez les patientes à bas risque génomique

⁸¹ Essai en ouvert à la fois pour l'investigateur et pour la patiente concernant l'estimation du risque génomique et clinique et les traitements reçus ; des taux importants de violation majeure au protocole de 20 % dans le groupe C Haut/G Bas ; le taux de sorties d'étude n'a pas été rapporté par les auteurs ; la nature des protocoles reçus de CTA et d'HTA n'ont pas été rapportés par les auteurs et était laissée à l'appréciation des investigateurs.

⁸² p = 0,055 en intention de traiter

(C-haut/G-bas) par rapport à celles à haut risque génomique (C-Bas/G-Haut) ; cet effet paradoxal n'étant pas souhaitable pour guider une désescalade par le test Mammaprint ;

- le seuil théorique de bon pronostic de 92 % défini par les auteurs est franchi dans la population ciblée l'évaluation (tumeurs RO+/HER2- pN0) devant une SSMD de 93,9 % IC 95% [90,6 – 96,1]. Les données comparatives sur la SSM à cinq ans n'ont pas été rapportées par les auteurs ;
- le bénéfice attendu de la désescalade sur le nombre d'EIG évités et l'amélioration de la qualité de vie n'a pas été rapporté par les auteurs, ce qui ne permet pas d'apprécier la balance bénéfice sécurité.

Les données publiées dans MINDACT ne permettent pas de s'assurer avec confiance que la perte de chances sur le contrôle du cancer en cas de désescalade de CTA basée sur un test Mammaprint à bas risque génomique soit cliniquement négligeable (< 3 %). Chez ces patientes, l'efficacité de la CTA est de plus conservée. Le manque de données pertinentes rapportées spécifiquement dans la population ciblée préalablement par la HAS (tumeurs RO+/HER2- pN0-1) ne permet de faire aucune conclusion quant à une balance bénéfice sécurité en situation de désescalade de CTA guidée par Mammaprint.

Oncotype Dx : étude TAILORx (110) (résultats cf. annexe 7)

L'objectif primaire de l'étude consistait à démontrer l'absence d'intérêt de la CTA chez des patientes à risque génomique intermédiaire (RS⁸³ entre 11 - 25 selon Oncotype Dx) par une analyse de non infériorité avec ou sans CTA. Les auteurs ont atteint l'objectif qu'ils s'étaient fixés. En intention de traiter, l'efficacité relative de la CTA était modeste avec un *hazard ratio* HR = 1,08 ; IC 95 % [0,94 - 1,24] avec un gain absolu de seulement 0,3 % sur la SSM à cinq ans⁸⁴ [IC 95 % non fourni] mais dont l'estimation de la borne supérieure serait proche de 1,5 % (118). Des constatations proches sont faites également en analyse selon le traitement reçu avec un HR 1,14 ; IC 95 % [0,99 - 1,31], borne supérieure très proche du seuil de non-infériorité [1,32], mais avec un gain absolu maximal de CTA probablement ≤ 3 % en raison du bon pronostic des patientes et d'une efficacité observée de CTA très modeste (118). Il est à signaler toutefois que de nombreuses violations majeures du protocole⁸⁵ ont été identifiées entre les bras randomisés (notamment dans le bras C sous CTA) au regard des exigences minimales de qualité d'une analyse de non infériorité réalisée en ouvert (119). De plus, l'absence de données rapportées de sécurité (EIG) et de qualité de vie à plus ou moins long terme ne permet pas d'estimer en confiance la balance bénéfice sécurité de la nouvelle stratégie.

La détermination de l'utilité clinique du test nécessitant une appréciation de ses résultats dans leur globalité, associant pour ce faire les résultats des deux autres sous-groupes de l'étude (RS≤10 et RS≥26). Dans le bras A (RS≤10), une désescalade de CTA était systématiquement décidée (patientes à bas risque génomique sous HTA exclusive) alors que dans le bras D (RS≥26), les patientes étaient systématiquement mises sous CTA (patientes à haut risque génomique).

Une population incluse de bon pronostic

Le profil de risque clinique des patientes de TAILORx était similaire à celui de la sous population d'intérêt des patientes pN0 à tumeur RO+/HER2- de l'étude MINDACT (environ 30 % de patientes à haut risque clinique dans les deux cas selon l'algorithme AOL modifié).

⁸³ Score génomique d'Oncotype Dx (*recurrence score*)

⁸⁴ 93,1 % sous CTA +HTA (bras C) versus 92,8 % sous HTA exclusive (bras B)

⁸⁵ Exclusion initiale de 2,8 % des patientes (bras B/C), **21,6 % de violations majeures au protocole dans le bras C** comportant 18,4 % de CTA non reçues pour 14 % de sorties prématurées, 7 % de violations majeures au protocole dans le bras B comportant 5,4 % d'HTA exclusive non reçues pour 11% de sorties prématurées.

Comme pour MINDACT, cet élément montre que le recrutement des patientes de TAILORx a été volontairement large et peu ciblé, générant une population hétérogène globalement de bon pronostic (le recrutement supplémentaire de 57 % de patientes dans les bras randomisés soutient cette surestimation initiale du risque réel des patientes par les auteurs). Toutefois, si dans MINDACT, toutes les patientes candidates à une désescalade de CTA, selon le test MammaPrint, (si faible risque génomique) étaient à risque clinique élevé de récurrence, seulement 25 % des patientes l'étaient dans TAILORx (patientes à $RS \leq 25$). Cette population candidate à une désescalade de CTA sur la base d'un risque génomique faible à intermédiaire était donc de meilleur pronostic que celle de MINDACT.

Parmi les patientes candidates, selon Oncotype Dx, à une désescalade de CTA ($RS \leq 25$), peu disposaient d'un critère majeur de mauvais pronostic :

- âge < 40 ans : 4,5 % ;
- pT2 > 3 cm (tumeur volumineuse) : 5 % ;
- envahissement ganglionnaire (critère d'inclusion de l'étude) : 0 % ;
- tumeurs de grade 3 : 12 % ;
- embolies vasculaires : non rapporté.

En outre, cette population candidate à une désescalade sur la génomique ($RS \leq 25$) disposait également de critères pronostiques favorables :

- tumeur infracentimétrique (pT1b) : 13 % (en dehors de la population préalablement ciblée par la HAS) ;
- tumeurs de grade 1 : 30 %
- tumeur « doublement » hormonosensible (RO+/RP+) : 91 - 93 %.

Au vu des critères d'éligibilité de l'étude, cette population ne pouvait représenter que partiellement la population préalablement ciblée par la HAS et susceptible de recevoir de la CTA sur une indication préalable (Q3). Les tumeurs pT1b étaient éligibles et aucun critère de mauvais pronostic n'était requis en cas de tumeur plus volumineuse (pT1c-pT2). L'analyse des critères de bon pronostic pour les tumeurs RO+/HER2- pN0 (notamment pT1b : 13 % et pT1c/2 grade 1 : 30 %) montre que plus de **40 % des patientes candidates à une désescalade de CTA** selon Oncotype Dx ($RS \leq 25$) **n'étaient pas ou peu susceptibles d'en recevoir** préalablement en contexte français.

De nouveaux seuils de décision thérapeutique basés sur des différences pronostiques observées à risque majeur de biais de sélection et d'attrition

Les auteurs signalent que les seuils de décision thérapeutique utilisés pour Oncotype Dx **depuis plus de 13 ans** ($RS < 18$, $RS 18 - 30$, $RS \geq 31$) ne doivent plus être les mêmes depuis les résultats de TAILORx (décision de CTA si $RS \geq 26$ pour les patientes de moins de 50 ans ou $RS \geq 16$ pour les patientes de plus de 50 ans).

Leurs conclusions se sont basées sur les **différences pronostiques rapportées** entre sous-groupes à risque génomique différents à cinq ans et neuf ans en population globale et chez les patientes de moins de 50 ans. Le schéma de l'étude TAILORx permettait principalement d'envisager un nouveau seuil de décision thérapeutique $RS \geq 26$ pour administrer la CTA une fois confirmée l'absence de son intérêt chez les patientes à risque génomique intermédiaire ($RS 11 - 25$). Le seuil de décision pour un $RS \geq 16$ chez les patientes de moins de 50 ans a été établi sur des analyses de sous-groupe non prévues au protocole.

Cependant :

- **aucun test statistique de supériorité** n'a été rapporté par les auteurs pour valider la plausibilité statistique des différences observées ;
- **aucune différence pronostique observée à neuf ans** ne pouvait être recevable pour l'ensemble des quatre bras de l'étude en raison d'un **risque majeur d'imprécision** avec des

données trop immatures ou manquantes⁸⁶. Ces données à neuf ans ne peuvent être retenues pour cette raison (120).

Les **données à cinq ans** pour les patientes RS ≤ 25 (Bras A, B, C) sont recevables en termes de précision et de maturité avec environ 80 % de patientes disposant de cette durée de suivi. La différence pronostique observée à cinq ans entre les patientes à risque génomique faible (bras A) et intermédiaire (bras B) sous HTA exclusive était **faible** (≤ 2 %) avec un pronostic excellent pour les deux groupes.

A contrario, la différence pronostique observée à cinq ans entre les patientes du bras C (RS 11 - 25) et celle du bras D (RS ≥ 26) sous CTA était **importante** (5 à 6 %). **Les auteurs recourent à cette différence pour valider le nouveau seuil de décision thérapeutique d'Oncotype Dx (RS ≥ 26) sans discuter de sa validité dans la publication.**

Cependant :

Le pronostic global des 1 737 patientes à RS ≥ 26 de TAILORx pourrait être très différent du pronostic rapporté à cinq ans des patientes du bras D pour les raisons suivantes :

- le bras D était une cohorte de suivi **non exhaustive à déclaration volontaire et non financée** par le promoteur de TAILORx comme pour les bras A, B et C (*cf.* supplément S5 de la publication) ;
- cette cohorte avait un **risque de biais de sélection majeur** *a fortiori* pour une **étude en ouvert** : 20 % des patientes RS ≥ 26 de l'étude ont été exclues dès le départ sans justification des auteurs et sans description de leur profil clinico-pathologique par rapport aux patientes analysées ;
- cette cohorte avait un taux important de violation au protocole : 25 % ;
- cette cohorte avait un **risque de biais d'attrition majeur** : selon la figure S4 en supplément de la publication, le faible nombre de patientes à cinq ans de suivi indiquait un nombre excessif et déséquilibré de données manquantes⁸⁷ par rapport aux autres bras (RS ≤ 25) pour seulement 50 sorties d'étude rapportées (« *flow chart* » de la publication). Avec un suivi clinique minimal théorique pourtant proche de sept ans⁸⁸, il existait donc à cinq ans **environ 50 % de données manquantes censurées**⁸⁹ à haut risque de biais (121) ;
- En cas de non prise en compte des **données censurées pour données manquantes**, les taux absolus d'événements cliniques (d'après le tableau S4 de la publication) restaient du même ordre entre d'une part les patientes sous CTA du bras D (13,6 %) et celles du bras C (12,1 %) et d'autre part celles sous HTA exclusive du bras B (12,8 %) et celles du bras A (11,4 %).

Pour déterminer le **seuil de décision thérapeutique RS ≥ 16 en cas d'âge de 50 ans ou moins**, ce sont des **analyses exploratoires et non prévues au protocole** dans des sous-groupes de relativement faible effectif⁹⁰ qui ont été retenues (Figure S11 & S13 de la publication). Cette conclusion se fondait sur la **significativité statistique** en termes d'efficacité relative de la CTA dans certains de ces sous-groupes et non sur la **significativité d'un test d'interaction statistique** entre ceux-ci et les autres, condition normalement nécessaire à la validation d'une valeur prédictive pour un test (122). Cette démarche exploratoire *a posteriori* a permis aux auteurs de proposer cet autre seuil (RS ≥ 16) aux patientes les plus jeunes. Cette démarche scientifique (seuils *post hoc*) n'est toutefois pas recommandée pour proposer une nouvelle prise en charge dans le cadre

⁸⁶ Au cours d'une étude, une proportion restante de patients **de moins de 20 %** à une date d'analyse donnée est généralement considérée comme insuffisante pour rapporter un résultat clinique (données trop immatures ou n'étant plus informatives en raison de données censurées importantes) (120). Or, pour estimer la probabilité de survie sans maladie à neuf ans, 13 à 15 % des patientes RS ≤ 25 et seulement 4,5 % des patientes RS ≥ 26 de l'étude (n=77) restaient encore dans TAILORx

⁸⁷ Pour l'estimation de la probabilité de survie sans maladie à cinq ans, seulement 36 % des 1 737 patientes RS ≥ 26 étaient encore suivies dans TAILORx

⁸⁸ Délai estimé entre la dernière patiente incluse en 2010 et le gel de la base de données en mars 2018

⁸⁹ Probabilité de survie sans maladie de 88 % pour 36 % encore à risque et sans événement clinique à cinq ans

⁹⁰ RS entre 21 et 25 (< 50 ans) : **492 patientes** et RS entre 16 et 20 (< 50 ans) : **923 patientes**

d'un essai thérapeutique d'oncologie et risque de complexifier d'autant les décisions des professionnels (123).

Il aurait été préférable que soient plutôt fournies les **différences pronostiques à cinq ans** dans des populations plus proches du **contexte français**, entre d'une part les sous-groupes RS ≤ 10 (bras A) et RS 11 - 25 (bras B) sous HTA exclusive et d'autre part entre les sous-groupes RS 11 - 25 (bras C) et RS ≥ 26 (bras D) sous CTA ; c'est à dire :

- chez toutes les patientes réellement susceptibles de recevoir une CTA (sans distinction), chez celles en incertitude décisionnelle (Q2) et chez celles ayant une indication préalable de CTA (Q3) ;
- chez les patientes à haut risque clinique (C-haut) selon l'algorithme AOL modifié.

Une valeur ajoutée pronostique et prédictive non démontrée par rapport aux CCPT

Les auteurs n'ont pas exploré la valeur ajoutée d'Oncotype Dx par rapport aux critères de mauvais pronostic. Ceux-ci étaient :

- beaucoup plus représentés chez les patientes RS ≥ 26 (**50 % grade 3**, 30 % RP négatif, 33 % de pT2 pour **7 % de grade 1**) que chez celles des bras A, B et C (12 % de grade 3 et 30 % de grade 1) ;
- les patientes RS ≥ 26 étaient deux fois plus à haut risque clinique (57 %) que les patientes des autres groupes (22 à 26 %) selon AOL modifié.

La **valeur prédictive** d'Oncotype Dx n'était ni alléguée par les auteurs, ni démontrée au regard des données de TAILORx⁹¹. Outre l'usage d'Oncotype Dx, **le jeune âge des patientes** (CCPT bien connu et validé) semblait être un critère prédictif majeur d'efficacité de la CTA dans TAILORx.

En raison d'une population différente et inappropriée (excellent pronostic sous HTA exclusive), la démonstration de l'utilité clinique d'Oncotype Dx utilisée pour décider d'une désescalade de CTA dans cette population ne peut être établie par les données de TAILORx. Ces données rassurantes mais dans une population globalement à faible risque de récurrence (1 à 3 % de récurrence à cinq ans) ne peuvent être transposées à une population avec de réelles indications de CTA sur des critères de mauvais pronostic. Avec ou sans SG, ces données suggèrent que la CTA pourrait ne pas apporter de bénéfice suffisant (< 3 %) chez la plupart des patientes de TAILORx. Les données spécifiques chez les patientes avec une indication de CTA (Q3) ou en incertitude décisionnelle (Q2) selon les critères de la HAS n'ont pas été fournies par les auteurs. Les nouveaux seuils de décision thérapeutique proposés par les auteurs (RS ≥ 26 si âge > 50 ans et RS ≥ 16 si âge < 50 ans) ne peuvent être validés au regard du manque de fiabilité et de précision d'une part du pronostic rapporté dans l'échantillon de patientes du bras D (RS ≥ 26) et d'autre part des analyses de sous-groupe réalisées sur quelques centaines de patientes. Ces nouveaux seuils stratifiés sur l'âge de 50 ans ont été fixés *a posteriori* sur la base d'analyses exploratoires qui demanderaient à être confirmées, *a fortiori* dans une population plus pertinente. Avec la modification et la complexification des seuils de décision avec TAILORx, ces changements significatifs pourraient continuer d'alimenter la confusion et l'hétérogénéité de pratiques parmi les cliniciens. La valeur ajoutée d'Oncotype Dx par rapport aux CCPT n'a pas été démontrée, sa valeur prédictive d'efficacité non plus.

⁹¹ Résultats non évocateurs en population globale (RS 11 - 25), différence de compliance à la CTA entre les sous-catégories bases et plus élevées de RS dans le groupe RS 11-25 (cf. supplément 6^E) sans analyse rapportée en traitement reçu ou en per protocole, pas de test d'interaction statistique entre les sous-groupes, design de l'étude non adapté pour ce type de démonstration (à la différence de l'étude RXPONDER)

Oncotype Dx : analyses ancillaires en marge de l'essai PLAN B (111) (cf. annexe 8)

En cours de recrutement de l'essai thérapeutique PLAN B⁹², un amendement au protocole a autorisé que certaines patientes ayant une tumeur RO+/HER2- pN0-1 à $RS \leq 11$ (et non $RS \leq 10$ comme TAILORx) ne soient pas retenues dans l'essai principal mais soient laissées sous HTA exclusive et suivies prospectivement en marge de celui-ci.

De manière générale, le **schéma de cette étude ancillaire en ouvert** ainsi que le **caractère très incomplet des données rapportées** par les auteurs ne permettent pas d'obtenir des preuves directes d'utilité clinique d'Oncotype Dx. Des incertitudes majeures demeurent sur les caractéristiques clinico-pathologiques des 348 patientes $RS \leq 11$ laissées sous HTA exclusive (description non spécifiquement rapportée). Les résultats n'étaient donc pas transposables à la pratique. En outre, 56 patientes $RS \leq 11$ ($n = 56$) n'avaient pas été retenues dans cette cohorte et ont reçu à la place une CTA sans explication ou justification des auteurs (risque de biais de sélection). Le pronostic **à cinq ans** des patientes $RS \geq 12$ mises sous CTA n'a pas été non plus rapporté.

Le pronostic global des patientes à risque génomique faible ($RS \leq 11$) sous HTA exclusive était, comme dans TAILORx, globalement bon⁹³ tout en manquant de précision en raison d'un faible effectif initial ($n = 348$) pour peu d'évènements cliniques observés et un suivi relativement immature à cinq ans⁹⁴.

Aucune comparaison entre catégories de risque génomique n'était recevable en l'absence de description des CCPT pour des traitements reçus différents. Aucun ajustement statistique approprié ne permettait de quantifier la valeur ajoutée d'Oncotype Dx par rapport aux CCPT.

L'utilité clinique du test Oncotype Dx ne pouvait être établie au vu du schéma d'étude proposé et des données cliniques très incomplètes fournies par les auteurs en marge de l'essai PLAN B. La sécurité de la désescalade de CTA, réalisée dans le cadre de cette étude à faible effectif, ne pouvait être transposable à la pratique en raison d'une population sous HTA exclusive aux caractéristiques clinico-pathologiques non décrites comparée à d'autres groupes très différents (sans randomisation ni ajustement statistique). La valeur ajoutée d'Oncotype Dx n'était pas démontrée par les auteurs.

3.2.5 Question 4 : les décisions de CTA sont-elles concordantes entre les différentes SG ?

► Objectif

Évaluer le degré de concordance décisionnelle de CTA entre les différentes SG.

► Sélection bibliographique

La sélection a permis d'identifier une seule étude de concordance décisionnelle prospective et comparative entre les SG Oncotype Dx, Mammaprint et Prosigna (OPTIMA, ISRCTN 42400492) (112).

Aucune étude prospective et comparative répondant à la question 4 n'a été identifiée avec Endopredict.

⁹² Essai de non infériorité comparant l'efficacité de deux protocoles de CTA différents

⁹³ Survie sans maladie à cinq ans : 94,2 % ; IC 95 % [91,2 – 97,3]

⁹⁴ 30 % des patientes ont atteint un suivi de cinq ans

► Principaux résultats de l'étude OPTIMA Prelim (112) (cf. annexe 9)

Ces résultats préliminaires de l'essai OPTIMA (en cours de réalisation) apportent des données prospectives et comparatives de concordance décisionnelle dans une **population à plus haut risque clinique**⁹⁵ selon AOL modifié.

Une comparaison des décisions de CTA entre trois SG différentes a été réalisée entre Oncotype Dx, Prosigna et MammaPrint.

Cette comparaison entre les trois tests a été possible pour 96,5 % des tumeurs disponibles (302/313 tumeurs).

La concordance décisionnelle était imparfaite de niveau **faible à modéré** selon les SG (IC à 95 % du score *kappa* compris entre 0,30 et 0,63 ; cf. annexe 9). **Entre 22 et 26 % des décisions se révélaient discordantes** dans la configuration de la présente évaluation « décision de CTA en cas de haut risque génomique ». Une désescalade de CTA a été plus fréquemment observée avec Oncotype Dx (seuil de décision $RS \geq 26$) qu'avec les deux autres SG.

► Conclusion

Une divergence de décision de CTA entre SG était retrouvée chez environ 1/4 des patientes. Cet élément soulève la question de leur interchangeabilité et du risque d'hétérogénéité de pratiques entre les différents centres de soins selon le choix de la SG et le seuil de décision thérapeutique qui auront été retenus.

3.2.6 Conclusion générale de la littérature

Concernant les situations d'incertitude décisionnelle (Q2) : aucune SG n'a été évaluée dans le cadre d'une étude d'impact et d'utilité cliniques en association avec les CCPT au sein d'un modèle de prédiction du risque de récurrence chez des patientes appartenant à la population ciblée préalablement pour l'évaluation.

Concernant les tests de deuxième génération (Prosigna/Endopredict) : aucune étude prospective et comparative évaluant l'impact et l'utilité cliniques de celles-ci n'a pu être identifiée. Leur utilité clinique ne peut donc être démontrée au regard de la littérature disponible encore immature.

Concernant le test de première génération Oncotype Dx : Cette SG a été principalement évaluée dans l'indication d'aide à une désescalade de CTA basée sur un score génomique rassurant ($RS \leq 18$ puis $RS \leq 25$) pour des patientes à tumeur RO+/HER2- de stade précoce avec une indication préalable de CTA sur la base des CCPT. Il a été retrouvé deux études prospectives disposant d'un suivi clinique (TAILORx et analyses ancillaires en marge de PLAN B). Les résultats de l'essai TAILORx ont été publiés très récemment (juin 2018). Les données disponibles ne permettent pas de déterminer avec confiance l'utilité clinique et la sécurité du test Oncotype Dx dans la population préalablement ciblée pour l'évaluation avec une indication de CTA sur la base de critères traditionnels de mauvais pronostic (Q3). La population incluse était large, hétérogène et de meilleur pronostic que celui attendu. Selon l'algorithme AOL modifié, environ 70 % des patientes de TAILORx étaient « à bas risque **clinique** de récurrence » comme 75 % de celles appartenant aux sous-groupes de patientes can-

⁹⁵ Patientes > 40 ans avec une tumeur RO+/HER2- ; pT2 (> 3cm) pN0 et/ou pN+ macrométastatique

didates à une désescalade de CTA sur la base du test ($RS \leq 25$). Le pronostic rapporté à cinq ans sous HTA exclusive chez ces dernières était excellent (motivant le recrutement de 57 % patientes supplémentaires lors d'un amendement du protocole). Aucune conclusion sur le seuil de décision $RS \geq 26$ ne peut être établie avec confiance au vu de la **faible qualité et de l'imprécision des données rapportées à neuf ans** (données manquantes, données immatures) et des données rapportées dans la cohorte ($RS \geq 26$) à haut risque génomique (exclusion de 20 % des patientes, données manquantes). La **différence pronostique rapportée à cinq ans** entre les patientes à haut risque génomique ($RS \geq 26$) et les autres ($RS \leq 25$). Cette différence pronostique à cinq ans est exposée à un **risque majeur de biais de sélection et d'attrition conduisant à une non représentativité de l'échantillon à haut risque génomique**. De plus, le second seuil de décision thérapeutique ($RS \geq 16$) chez les patientes de moins de 50 ans ne peut être validé avec fiabilité car il est issu d'analyses multiples et non prévues au protocole (recherche exploratoire) concernant des tailles d'efficacité relative de CTA sur de relativement faibles effectifs. La valeur ajoutée pronostique du test Oncotype Dx et sa valeur prédictive n'ont pas pu être démontrées par les auteurs. La balance bénéfice sécurité ne peut être estimée en l'absence de données de sécurité publiées (EIG). Aucune conclusion pour la pratique ne peut être formulée sur la base des résultats de l'étude ancillaire à PLAN B en raison de données très incomplètes et peu informatives à cinq ans.

Concernant le test de première génération Mammaprint : L'impact clinique de cette SG a été évalué par l'étude MINDACT conçue il y a plus de dix ans dans les deux indications : i) capacité à remplacer la stratégie habituelle dans une population large et hétérogène afin d'optimiser la décision (Q1), ii) capacité à décider d'une désescalade de CTA en cas d'indication préalable sur des critères de mauvais pronostic (Q3). Au vu des données publiées, la stratégie basée sur Mammaprint n'a pas démontré sa capacité à améliorer le devenir clinique des patientes de la population préalablement ciblée pour l'évaluation (tumeurs RO+/HER2- pN0-1 de stade précoce sans distinction). Le test Mammaprint n'a pas fait non plus la preuve directe de son utilité clinique en situation de désescalade de CTA. Une perte de chances a été statistiquement et cliniquement constatée sur la survie sans maladie à cinq ans dans une population plus large et moins ciblée que la population préalable de l'évaluation sans qu'un gain de sécurité (EIG) ou d'amélioration de qualité de vie ait pu être rapporté (Q3). Aucune donnée cliniquement pertinente pour une situation de désescalade de CTA n'a été rapportée dans la population ciblée pour l'évaluation (tumeurs RO+/HER2- pN0-1).

3.3 Position du groupe de travail

Cette partie synthétise la position argumentée et multidisciplinaire des experts réunis physiquement à la HAS le **02 mai 2018** (discussion sur le protocole d'évaluation préparé par la HAS) et le **02 juillet 2018** (discussion sur l'analyse des données de la littérature préparée par la HAS) dont les comptes rendus validés figurent *in extenso* respectivement en annexe 10 et en annexe 11.

3.3.1 Préambule des experts

Les experts ont d'abord voulu souligner la bonne couverture géographique et le savoir-faire des professionnels français impliqués dans l'analyse des tumeurs du sein (anatomopathologistes, biologistes médicaux) ainsi que la rigueur du contrôle qualité des laboratoires qui reste une spécificité française.

Selon les experts du GT, les SG n'ont pas vocation à se substituer aux CCPT pour décider d'une CTA mais à les compléter dans certains cas en vue d'optimiser la décision finale. À ce jour, le besoin des professionnels dans le cancer du sein de stade précoce est d'améliorer la décision de CTA dans les situations d'incertitude décisionnelle.

► Considérations générales

Les experts reconnaissent la valeur pronostique intrinsèque des SG eu égard, selon eux, aux nombreuses études de validation clinique dans la littérature mais trouvent néanmoins difficiles d'identifier leur valeur ajoutée par rapport aux autres CCPT.

Le groupe d'experts demande l'obtention de preuves directes de l'amélioration de la santé des patientes (impact clinique) permettant de démontrer l'utilité clinique des SG ; aucun consensus n'ayant été trouvé en ce qui concerne le mécanisme sous-jacent à privilégier (améliorer l'exactitude pronostique ou bien sélectionner les patientes avec une efficacité anticipée de la CTA suffisamment importante).

Le GT a particulièrement souligné que l'utilisation d'un biomarqueur pour décider d'une désescalade thérapeutique en cas d'indication préalable de CTA expose les patientes à des déviations de pratique graves pour cette maladie curable. La validation d'une indication nécessite d'obtenir des données cliniques fiables, transposables à la pratique et rassurantes. Ce dernier élément a été particulièrement discuté par les experts compte tenu de messages d'information ayant circulé récemment dans la presse grand public qui annonceraient « l'inutilité de la CTA dans 70 % des cas de cancer du sein » depuis les résultats de TAILORx.

► Encadrement de la prescription de SG

Selon eux, l'encadrement nécessaire devrait se concevoir par un examen préalable et collégial en RCP une fois les CCPT obtenus. La prescription effective d'une SG devrait se faire après une information de la patiente de son risque de récurrence suivie du recueil de ses préférences en matière thérapeutique et de son accord dans le cadre d'une décision médicale partagée qui favorisera sa compliance au résultat du test à venir.

3.3.2 Remarques sur le protocole d'évaluation proposé par la HAS

Au sein de la population ciblée pour l'évaluation (tumeurs RO+/HER2- pT1c-2, pN0-1), les experts ont confirmé qu'il existe deux situations d'intérêt distinctes pour le clinicien :

- les cas d'incertitude décisionnelle (critères intermédiaire/équivoques, discordance clinico-histopronostique) ;
- les cas d'indication préalable de CTA (sur des critères de mauvais pronostic).

3.3.3 Critères de mauvais pronostic

Selon le GT, les critères majeurs de mauvais pronostic sont : le jeune âge (< 35/40 ans), les tumeurs pT2 (T > 2-3 cm), le grade 3 histopronostique, le statut pN1 macrométastatique, la présence d'emboles lymphovasculaires.

Selon eux, une indication préalable de CTA serait quasi-systématique en contexte français en présence d'au moins deux de ces critères de mauvais pronostic dans la population ciblée de l'évaluation. La présence d'un seul critère comme un grade 3⁹⁶ ou la présence de macrométastases ganglionnaires (pN1) dans cette population pT1c-pT2 pourrait conduire certains centres à administrer fréquemment une CTA.

3.3.4 Place du Ki67

Ce marqueur de prolifération est largement utilisé dans la pratique en raison de sa valeur pronostique supposée et d'un financement à la CCAM non spécifique. Cependant, son utilité clinique pour décider d'une CTA n'est pas communément admise en raison de nombreuses limites telles que sa reproductibilité, ses résultats équivoques fréquents, ses seuils décisionnels peu consensuels. Ce marqueur est utilisé dans le cadre de plusieurs modèles de prédiction du risque en cours de vali-

⁹⁶ Le devenir des patientes à tumeur RO+/HER2- pT1b pN0 de grade 3 (hors de la population ciblée pour l'évaluation) mériterait d'être exploré par une étude pronostique française sur des données récentes et de bonne qualité

dition (IHC4 +C, PREDICT v2.0) mais pas comme un critère majeur de mauvais pronostic pour décider d'une CTA.

► Critères de jugement des études

Si le pronostic vital des patientes dans le cancer du sein s'estime par la survie globale ou l'apparition d'une métastase à distance, l'impact thérapeutique d'un traitement en situation adjuvante quant à lui s'évalue sur la survie sans maladie (ou sans rechute) et sa toxicité induite (effets indésirables graves aigus et différés). La survie sans maladie est le critère primaire des essais thérapeutiques en contexte adjuvant pour l'EMA et la FDA (96, 100) mais également pour les experts internationaux (73, 94, 95, 123). La qualité de vie reste un critère secondaire (en situation adjuvante) et le suivi clinique devrait être de cinq ans minimum.

► Modèle de prédiction, seuil de risque et bénéfice anticipé

Les modèles de prédiction du risque seraient peu utilisés en France pour décider d'une CTA. En pratique, un risque de décès à dix ans de plus de 10 % pourrait justifier selon certains une CTA. Un bénéfice à cinq ans sur les récidives < 3 % ne justifierait plus la CTA en raison de sa toxicité. Un bénéfice > 5 % la justifierait clairement. Entre 3 et 5 %, cette décision devrait être partagée avec la patiente.

3.3.5 Remarques sur l'analyse critique des données de la littérature

Certains experts⁹⁷ ont souligné la difficulté d'appropriation en première lecture de ces publications. Concernant les résultats rapportés dans les publications, les choix de notification arbitrés par les auteurs pouvaient être perçus comme perfectibles et incomplets au vu de ceux attendus.

Les experts ont été surpris des populations incluses dans ces études⁹⁸, de la faible qualité des données rapportées pour des essais thérapeutiques (écarts au protocole, perdus de vue et données manquantes, absence d'insu) et des moyens logistiques et financiers mis en œuvre (plusieurs milliers de patientes) pour démontrer des éléments qui leur semblaient pourtant connus (désescalader une CTA chez une patiente peu susceptible d'en recevoir sur la base des CCPT).

Ainsi, selon eux, les résultats des deux grands essais MINDACT et TAILORx⁹⁹ ne permettaient pas de démontrer l'utilité clinique des tests Mammaprint et d'Oncotype Dx avec l'objectif de modifier les pratiques dans les populations ciblées pour l'évaluation.

Les membres du GT constatent qu'il n'existe pas de littérature disponible permettant d'apporter des données d'utilité clinique pour les tests Endopredict et Prosigna. De même, aucune étude prospective d'impact clinique n'a évalué l'utilité clinique d'intégrer un score génomique au sein d'un modèle de prédiction du risque validé.

Un sujet a suscité également l'inquiétude des experts. Il concerne le fait que ces tests conduisent à des discordances décisionnelles (dans 20 à 25 % des cas) dont la fréquence ne peut être négligée du point de vue des professionnels. Le GT estime que ce résultat pourrait être encore accentué dans une population en incertitude décisionnelle (principale cible de patientes pour les experts).

3.3.6 Propositions et perspectives formulées par les experts

Concernant la prescription des SG, les experts se positionnent à ce jour pour :

- leur non indication en cas de tumeur non luminale (HER2+, RO-, triple négatifs) ou chez des patientes pas ou peu susceptibles de recevoir de la CTA en contexte français (pronostic excellent) ;

⁹⁷ Dont des méthodologistes spécialisés en oncologie

⁹⁸ Spectre d'éligibilité trop large et mal défini par rapport aux populations ciblées pour l'évaluation

⁹⁹ Et plus accessoirement l'analyse ancillaire en marge de PLAN B

- leur contre-indication en cas de tumeur localement avancée (pT3-4 et/ou pN2-3) ou précoce mais possédant des critères de mauvais pronostic en faveur d'une indication de CTA.

Les experts souhaitent que les efforts de recherche se focalisent sur les patientes en **situation d'incertitude décisionnelle** (zone grise du clinicien) pour lesquelles peu de données cliniques sont réellement disponibles. Selon eux, la distinction entre cette zone d'incertitude et les zones relevant d'une indication ou d'une non indication de CTA dans le cancer de stade précoce mériterait d'être éclaircie afin de lutter contre l'hétérogénéité des prescriptions entre centres. A cet effet, l'élaboration d'un document commun de type recommandation nationale de bonne pratique a été proposée¹⁰⁰.

Deux profils distincts de patientes conduisent actuellement les cliniciens vers une situation d'incertitude décisionnelle sur la base des CCPT :

- cas équivoques à degré de prolifération intermédiaire ;
- cas de discordance clinico-histo-pronostique complexe.

Le GT recommande en priorité d'évaluer la valeur ajoutée des SG dans une population bien définie, homogène de cas équivoques à prolifération intermédiaire. Cette population en incertitude est très proche de celle proposée par le groupe de Saint Paul de Vence en 2015 : **Adénocarcinome (préférentiellement canalaire)¹⁰¹ RO+/HER2- pT1c-T2 ; pN0-1mi et de grade 2.**

Les cas plus complexes liés à une discordance entre les nombreux CCPT ne représentent pas une population de choix pour évaluer prioritairement leur valeur ajoutée en raison d'une hétérogénéité de situations rencontrées. Pour ces cas complexes, la réponse préalable à apporter serait le recours à un modèle de prédiction estimant le risque individuel des patientes (en termes de récurrence et de bénéfice de CTA). Ce modèle, une fois validé en contexte français, pourrait intégrer un score génomique qui aura apporté la preuve de sa valeur ajoutée (Q2).

Concernant l'évaluation de la valeur ajoutée des tests Endopredict et Prosigna de développement plus récent, le groupe propose une analyse attentive des données analytiques déjà recueillies¹⁰² à travers le financement RIHN 2016 - 2018 et un renforcement des exigences de cette collecte¹⁰³ dans la population en incertitude à « profil biologique intermédiaire » si celle-ci est prolongée (profil clinico-pathologique et suivi clinique).

3.3.7 Synthèse

De l'avis des experts, si une valeur pronostique intrinsèque n'est pas à remettre en cause au vu des études de validation, l'absence de donnée clinique, à ce jour, quantifiant concrètement la valeur ajoutée et l'utilité clinique des SG (par rapport aux CCPT) pour optimiser la décision de CTA, reste problématique. De plus, **une SG ne devrait pas se substituer aux CCPT pour décider d'une CTA**. Selon eux, leur valeur prédictive n'a pas été, à ce jour, démontrée par les études prospectives d'impact clinique à disposition.

Avec les deux grands essais MINDACT et TAILORx, **les SG de première génération** (Mammaprint et Oncotype Dx) seraient arrivées selon les experts à un stade final de leur développement en population pN0. A ce stade, les données cliniques disponibles n'ont pas fait la preuve directe de leur utilité clinique dans cette population¹⁰⁴.

¹⁰⁰ Le pilotage du projet par l'INCa a été proposé

¹⁰¹ Type histologique le plus fréquent (80 %) et le plus agressif par rapport aux adénocarcinomes lobulaires et aux tumeurs non adénocarcinomeuses (tubuleuses...)

¹⁰² Force des corrélations entre le score génomique et les critères biologiques non génomiques (grade, ki67...),

¹⁰³ Distribution/dispersion des valeurs de score génomique dans une population à profil biologique intermédiaire

¹⁰⁴ Des données d'impact clinique en population N+ sont en attente pour Oncotype Dx (essai RXPONDER et WSG ADAPT R0+/HER2-)

Les SG de deuxième génération (Prosigna et Endopredict), quant à elles, restent toujours en attente des premiers résultats d'utilité clinique¹⁰⁵. Les données disponibles à ce jour ne sont pas suffisantes.

Au total :

Les experts considèrent qu'il existe aujourd'hui un besoin d'encadrement de cette prescription de SG dans le cancer du sein de stade précoce (critères d'éligibilité plus restrictifs et mieux ciblés).

Le GT soutient l'absence de leur indication en cas :

- de statut tumoral immunohistochimique RO- et/ou HER2+ (triple négatifs, RO- ; HER2+) ;
- de patientes peu ou pas susceptibles de recevoir une CTA en contexte français (sans critère de mauvais pronostic).

Le GT soutient également que, dans l'attente des résultats des essais en cours en population N+ et/ou avec une indication établie de CTA, il existe une contre-indication à leur réalisation en cas :

- de tumeur localement avancée (pT3-4 et/ou pN2-3) ;
- d'indication préalable de CTA dans la population ciblée pour l'évaluation fortement évoquée par la présence d'au moins deux critères de mauvais pronostic, ou d'un profil biologique agressif (grade 3 histopronostic) ou d'une extension ganglionnaire macrométastatique (N+).

Le GT soutient le besoin et la poursuite d'une recherche pour les patientes en incertitude décisionnelle. Cette population pourrait être explorée dans le cadre d'un soutien financier de la recherche et de l'innovation (de type RIHN). Elle concernerait les patientes ciblées par l'évaluation, à la fois susceptibles de recevoir une CTA en contexte français, mais en situation d'incertitude décisionnelle où peu de preuves valident clairement l'intérêt clinique de ce traitement. Ces tumeurs devraient présenter en particulier un degré de **prolifération intermédiaire** (cas équivoques les plus simples et homogènes dans une optique de recherche) :

- adénocarcinome invasif (préférentiellement de type canalaire) ;
- RO+/HER2- ;
- pT1c-T2 ;
- pN0-N1mic ;
- de grade 2.

Pour les cas complexes d'incertitude décisionnelle en rapport avec une discordance clinico-histo-pronostique, le recours préalable à un modèle de prédiction du risque individuel des patientes (nomogramme) qui reste à valider en contexte français serait à privilégier. Ce modèle devrait intégrer le maximum de critères pronostiques et prédictifs indépendants les uns des autres (parmi l'âge, l'extension tumorale et le profil biologique de la tumeur) et être validé sur des données françaises récentes et de bonne qualité. Cet algorithme décisionnel permettrait d'explorer la valeur ajoutée des scores génomiques dans l'optique de les y intégrer si nécessaire.

Selon les experts, un passage préalable du dossier en RCP est impératif avant toute décision définitive de prescription. La patiente éligible devra être ensuite informée de façon claire, loyale et appropriée sur les faits établis de la science. Pour ces patientes en incertitude décisionnelle chez qui le bénéfice de la CTA peut être modeste, leur préférence en matière de choix thérapeutique et leur accord doivent être recueillis impérativement et pris en compte dans le cadre d'une décision médicale partagée.

¹⁰⁵ L'étude OPTIMA est en cours de réalisation pour Prosigna. Aucune étude de ce type n'est prévue à ce jour pour Endopredict et MammaPrint

Les experts soulignent le besoin d'harmoniser les prescriptions de CTA entre les centres d'oncologie français. Une clarification au niveau national des indications et des non indications évidentes ferait ressortir utilement la zone d'incertitude décisionnelle. À cet effet, ils proposent que soit élaboré un document commun de recommandations de bonne pratique pour les tumeurs RO+/HER2- de stade précoce. Ce document serait idéalement piloté par l'INCa avec l'aide des professionnels de santé, afin de réduire la part d'hétérogénéité des prescriptions de CTA en lien avec des référentiels régionaux plus ou moins divergents.

Dans tous les cas, l'évaluation du **risque de récurrence tumorale** (sous HTA exclusive) de ces patientes en incertitude décisionnelle pourrait être explorée par des études pronostiques sur des bases de données françaises et récentes de bonne qualité afin d'estimer le bénéfice envisageable de la CTA chez elles.

3.4 Position institutionnelle de l'INCa

Par la voix de son président, l'INCa a tenu à souligner par courriel la position institutionnelle *in extenso* suivante :

« [J]'ai plaisir à [vous] confirmer la grande qualité de ce rapport que nous tenons à saluer. L'analyse des données de la littérature et la position des experts du groupe de travail sont à la mesure des attentes exprimées par les collègues et les malades. Après leur validation finale, les conclusions nous semblent à même de définir de façon convaincante des situations cliniques précises pour lesquelles les SG seront contre-indiquées et d'autres pour lesquelles leur utilité clinique reste insuffisamment démontrée. Ce document met également en évidence une corrélation médiocre entre les résultats de ces différentes SG, ce qui peut conduire à des discordances non négligeables entre les décisions thérapeutiques ainsi induites. Il est regrettable que les recherches conduites ces dernières années pour développer et valider différentes SG, en particulier dans les phases précoces, se révèlent insuffisantes pour permettre aujourd'hui leur utilisation clinique. A vrai dire, la construction même de plusieurs de ces essais permettait de prévoir leurs limites, donc leur incapacité à emporter la conviction. La description précise des réponses à apporter figure en bonne place au nombre des grandes qualités de ce rapport. L'Institut National du Cancer prendra bien entendu acte des recommandations et conclusions finales, et l'appui de ce document sera précieux dans ses travaux à venir. Il nous semble essentiel que le Référentiel des Actes Innovants Hors Nomenclature (RIHN) soit mis en cohérence avec les conclusions de ce rapport : Si les tests évalués apparaissent d'ores et déjà inutiles dans certaines situations cliniques, le financement dérogatoire de ces actes mérite de contribuer de manière significative aux réponses qui manquent encore. »

3.5 Points de vue des parties prenantes

3.5.1 Préambule

Les parties prenantes (PP) ont été consultées afin de recueillir leurs points de vue sur la lisibilité et la clarté du rapport, sur leur appropriation du travail et les conséquences pratiques et organisationnelles qu'ils souhaitent faire remonter. Les réponses *in extenso* des organismes consultés sont rapportées en annexe 12.

Sur les **neuf organismes professionnels et associations de patients** sollicités, huit ont répondu¹⁰⁶. Sept au travers du questionnaire prévu à cet effet, parmi eux l'association de patients « **Vivre comme avant** », qui a précisé n'avoir aucune remarque sur le rapport provisoire. La **SFSPM**¹⁰⁷ qui a répondu sous format libre, a essentiellement apporté des remarques complémentaires sur certains éléments utiles aux lecteurs du rapport mais n'a pas formulé de point de vue sur son degré d'appropriation des conclusions et perspectives provisoires proposées par la HAS.

Dans l'ensemble, la majorité des PP, tant les organismes professionnels comme le **CNPP**¹⁰⁸, le **CNPBM**¹⁰⁹ et le **GFCO**¹¹⁰, que la **LNCC**¹¹¹ (seconde association de patients), a reconnu la qualité et la rigueur du rapport provisoire. Ils ont par ailleurs souligné la pertinence des questionnements cliniques soulevés (notamment la question d'évaluation n° 2) et le caractère factuel de l'analyse issue d'une recherche exhaustive de la littérature reposant sur une analyse critique méthodique.

Aucune des parties prenantes consultée n'a remis en cause les conclusions du rapport provisoire sur l'absence de démonstration de l'utilité clinique des SG dans les populations d'intérêt ciblées par l'évaluation de la HAS, **à l'exception de la SFMPP**¹¹². Selon le **CNPP** et le **GFCO**, l'absence actuelle de cette démonstration est « sans ambiguïté » au regard des données présentées dans le rapport.

3.5.2 Principaux éléments remontés

La synthèse¹¹³ des retours de la majorité des parties prenantes (à l'exception de la **SFMPP**) a permis de ressortir les éléments suivants :

► Rôle toujours essentiel des CCP traditionnels pour évaluer le pronostic d'un cancer du sein

Le **CNPP** rappelle le rôle essentiel des CCPT issus de l'examen anatomopathologique dans l'évaluation du pronostic d'un cancer du sein. Le **CNPBM** considère que dans le rapport les phénomènes biologiques impliqués dans les SG ne sont pas décrits, ce qui selon lui aurait permis de constater une « redondance avec les éléments phénotypiques [c'est-à-dire, les CCPT] utilisés en pratique courante pour apprécier le pronostic et les indications thérapeutiques dans le cancer du sein ».

► Complexité de la littérature disponible sur le sujet

A l'instar de la **SFSPM**, le **CNPBM** a souligné la difficulté de lecture des publications retenues, qui s'adressent selon lui à des « méthodologistes confirmés ».

► Difficulté à transposer les populations des grandes études au contexte français

La **SFSPM** a voulu souligner l'existence de « multiples raisons », qui font que l'utilisation des SG par les cliniciens soit différente entre les contextes français et nord-américain, en indiquant à titre d'exemple le « contrôle de qualité en anatomopathologie ». Sans se positionner, le **CNPP** appelle le rapport à mieux préciser les « particularités » françaises (par exemple : les référentiels, l'organisation des soins, les modalités de financements).

¹⁰⁶ Le CNP d'oncologie médicale n'a pas donné suite à notre sollicitation

¹⁰⁷ Société française de sénologie et de pathologie mammaire

¹⁰⁸ CNP de pathologie

¹⁰⁹ CNP de biologie médicale

¹¹⁰ Groupe francophone de cytogénétique oncologique

¹¹¹ Ligue nationale contre le cancer

¹¹² Société française de médecine prédictive et personnalisée

¹¹³ Les différentes remarques relevant d'un oubli, d'une insuffisance ou d'une imprécision du rapport provisoire ont été le plus souvent intégrées dans le rapport final sans faire l'objet d'un référencement détaillé dans cette partie (cf. annexe 12)

► Confirmation que la population d'intérêt prioritaire est celle en incertitude décisionnelle

Les sénologues de la **SFSPM** et le **CNPP** ont indiqué qu'au vu du recul et de l'expérience acquise, les deux grandes études d'impact clinique publiées (MINDACT/TAILORx) ne sont plus adaptées et ne ciblent pas « la vraie zone grise » de la population qui est restreinte aux patientes « en incertitude décisionnelle ». Selon eux, ces études conçues il y a plus de dix ans ne permettent pas de répondre aux questions posées aujourd'hui par les professionnels. Dans leur majorité, les organismes professionnels sont en accord avec cette population d'intérêt identifiée par l'évaluation.

Ils ont en outre indiqué qu'une définition plus précise et plus claire, selon des critères d'éligibilité plus « restrictifs » pour cette population en incertitude décisionnelle, devait être apportée dans les conclusions du rapport final. Ces demandes sont en opposition avec le point de vue affiché par la **SFMP** qui ne demande ou ne propose aucun encadrement particulier.

► Utilisation des SG dans un but pronostique exclusif

A l'instar de la **SFSPM** et du **GFCO**, le **CNPP** rappelle que les SG dans le cancer du sein de stade précoce « ont été conçues initialement pour une évaluation pronostique » et « non une utilisation prédictive » d'efficacité des CTA. Ces organismes insistent sur le fait, qu'à l'heure actuelle, les données d'utilité clinique ne permettent pas de concevoir une utilisation de ces tests autre que pronostique.

► Pertinence d'évaluer l'utilité clinique des SG

Le **CNPP**, le **GFCO** et le **CNPBM** confirment l'importance d'évaluer en priorité l'utilité clinique des SG pour évaluer leur intérêt. Toutefois, ils tiennent à souligner le nombre limité de publications retenues dans le rapport, situation qui pourrait être perçue par certains comme « une fragilité » potentielle. Certaines parties prenantes ont rappelé à cette occasion l'existence de trois grandes études observationnelles récemment publiées¹¹⁴ pouvant valider les performances pronostiques de différentes SG sur des cohortes nationales (par exemple : études SEER¹¹⁵, CLALIT¹¹⁶, DCBG¹¹⁷) ou encore d'une autre étude¹¹⁸ comparant rétrospectivement les performances de trois SG différentes sur un même échantillon de tumeurs.

¹¹⁴ La HAS a souhaité tirer expérience des problématiques rencontrées par l'INCa (34) et très récemment par l'INESS (Institut National d'excellence en santé et en services sociaux du Québec) (35). Ils ont rappelé l'extrême difficulté d'émettre des conclusions fiables en faveur de l'utilité clinique d'un biomarqueur avec des analyses rétrospectives issues de biobanques, quelles que soient leur constitution. La HAS a fait le choix de s'en tenir aux études comparatives d'impact clinique construites pour évaluer directement l'intérêt clinique des SG. Ces études prospectives et comparatives sont construites normalement autour d'une population pertinente selon un protocole préenregistré, des critères de jugement et un plan d'analyse statistique définis a priori susceptibles d'apporter des preuves directes. Ce choix a été fait dans l'optique de pouvoir modifier en toute sécurité la prise en charge d'une maladie grave et curable requérant un traitement lourd mais efficace (la CTA). Les grandes bases de données « en vie réelle », dont il faut faire à chaque fois la preuve de leur exhaustivité à des fins de recherche, génèrent principalement des analyses de sous-groupe pour des populations sélectionnées a posteriori plus éloignées de la cible idéale. Ces études ne sont pas le gold standard pour valider avec confiance une nouvelle prise en charge thérapeutique. Ces cohortes observationnelles peuvent toutefois avoir l'ambition de compléter l'information apportée par un essai randomisé qui aurait été concluant (maintien en vie réelle de l'utilité du test, conditions d'utilisation en vie réelle, extension à des sous-groupes peu représentés dans l'essai princeps, recueil d'événements indésirables rares ou à long terme, évaluation médico-économique) (124)

¹¹⁵ Roberts MC *et al.* *Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database.* *Breast Cancer Res Treat.* 2017 Jun;163(2):303-310.

¹¹⁶ Stemmer SM *et al.* *Clinical outcomes in ER+ HER2 -node-positive breast cancer patients who were treated according to the Recurrence Score results: evidence from a large prospectively designed registry.* *NPJ Breast Cancer.* 2017 Sep 8;3:32.

¹¹⁷ Lænkholm AV *et al.* *PAM50 Risk of Recurrence Score Predicts 10-Year Distant Recurrence in a Comprehensive Danish Cohort of Postmenopausal Women Allocated to 5 Years of Endocrine Therapy for Hormone Receptor-Positive Early Breast Cancer.* *J Clin Oncol.* 2018 Mar 10;36(8):735-740.

¹¹⁸ Les comparaisons rétrospectives entre Oncotype Dx, Prosigna et Endopredict proposées par Sestak *et al* en 2018 (125) s'appuyaient sur des paramètres sans signification évidente pour le clinicien (HR, rapport de vraisemblance, c-

► Le point de vue divergent de la SFMPP

A l'inverse du point de vue affiché par les autres organismes, la **SFMPP** considère que le rapport provisoire manque de clarté et d'objectivité et qu'il a une lecture non pertinente des données disponibles. Elle émet un avis réservé sur son analyse et ses conclusions.

Pour la **SFMPP**, le rapport provisoire n'est pas en ligne avec l'intérêt des malades, ni avec le rôle à venir de la médecine génomique. Selon cet organisme, les principaux écueils du rapport sont une mauvaise appropriation des objectifs et des caractéristiques des patientes dans les études retenues. Cet organisme reproche au rapport une forme de confiance excessive envers les CCPT pour décider d'une CTA.

La **SFMPP** « pense souhaitable d'intégrer rapidement en pratique les SG, dont certaines sont aujourd'hui validées pour certains sous-groupes de patientes [...] permettant ainsi d'éviter des effets secondaires et certaines dépenses inutiles ». Elle alerte notamment sur le risque « des réactions négatives¹¹⁹ possibles des associations de patientes, des sociétés savantes et de la communauté médicale, si le rapport reste en l'état » Selon elle, aucune « exception française » ne devrait avoir lieu qui pourrait placer les professionnels concernés devant « un décalage par rapport aux recommandations internationales, [...] [qui constituerait] une perte de chances pour les patientes, une réelle difficulté d'exercice pour le médecin qui a une obligation de moyen et [...] [qui créerait] une non équité d'accès aux soins ».

La **SFMPP** a rédigé un argumentaire libre qui revendique l'utilité clinique des SG pour la désescalade de la CTA, en particulier chez les patientes à risque « intermédiaire bas » qui, selon eux, sont en effet bien représentées dans les études analysées. En opposition aux autres PP, la **SFMPP** a voulu rappeler leur valeur prédictive et l'apport des données d'impact décisionnel pour démontrer la pertinence d'une désescalade guidée par une SG. Elle a également souligné leur intérêt pour harmoniser les prises en charges des cliniciens, selon eux trop hétérogènes, en raison d'une variabilité excessive entre les différents référentiels concernant le panel et les seuils décisionnels choisis pour les CCP. La **SFMPP** a notamment indiqué qu'il fallait « émettre des réserves » quant à « la rigueur du contrôle qualité » de certains CCP traditionnels comme le « grade [histopronostique] » et la « présence d'embolies vasculaires » au risque d'induire une « surévaluation » du pronostic.

Enfin, selon la **SFMPP**, les deux essais « de grande ampleur » et « indépendants de l'industrie » qui ont été analysés dans le rapport provisoire (TAILORx, MINDACT) possèdent comme principales atouts d'être « les plus grandes études prospectives randomisées jamais réalisées en cancérologie, et [d'être] publiées dans le New England Journal of Medicine ». Cet organisme ne discute toutefois ni des résultats publiés ni de la qualité des données rapportées.

3.5.3 Principales perspectives d'avenir soulevées

► Définition précise du champ d'utilisation d'une SG

L'ensemble des parties prenantes (à l'exception de la **SFMPP**) a fait remarquer qu'il était primordial que soient mieux définies et mieux encadrées l'usage des SG dans le but de les restreindre à une population en incertitude décisionnelle.

index) montrant un lien d'association concernant un évènement clinique peu pertinent en contexte adjuvant et à risque potentiel de biais (délai avant métastase). Une de ces SG (prosigna) avait été développée sur cet échantillon par le passé générant un risque de surestimation de ces performances par rapport aux autres tests. Les IC 95 % de ces paramètres se chevauchaient tous, sans qu'un test de supériorité statistique ne soit rapporté par les auteurs. Enfin, cette étude comparative, même ambitieuse, s'appuyait sur un échantillon tumoral R0+/HER2- pertinent mais ne représentant qu'environ 15 % de la population R0+/HER2- de l'essai d'origine ATAC (126). Dans ce contexte, aucune hiérarchisation entre les trois SG ne pouvait être proposée en l'absence de démonstration probante basée sur des critères cliniques robustes

¹¹⁹ Les deux associations de patients et les autres organismes professionnels n'ont pas eu ce genre de « réactions négatives » et ont souligné l'importance de clarifier et d'encadrer cette utilisation

► Réévaluer le périmètre, les objectifs, (et le cas échéant) l'existence du financement RIHN pour les SG

Les organismes professionnels s'interrogent tous, y compris la **SFMPP**, sur une remontée d'informations qui n'a clairement pas fonctionné de manière homogène entre les différentes SG mises sur le marché.

Ils soulignent globalement leur déception vis à vis des objectifs initiaux de recueil de données comparatives et cliniques pour les SG financées transitoirement dans le cadre du RIHN¹²⁰. En l'absence aujourd'hui de données cliniques suffisamment fiables, les organismes professionnels appellent à une réévaluation pour les SG de son périmètre et de ses objectifs, voire même de son existence pour la **SFMPP**, à court et moyen terme.

A l'instar du **CNPBM**, le **GFCO** estime que tout maintien de ces SG au financement RIHN, nécessiterait le cas échéant, de définir avec précision la population d'intérêt et d'établir des indications non ambiguës pour ces tests. Selon eux, une prescription uniquement après décision collégiale en RCP avec une obligation de recueil et de remontée de données cliniques en vie réelle est indispensable. Devant l'absence de données suffisamment fiables, le **CNPBM** indique que la prescription des SG pourrait être financée également par des programmes spécifiques de recherche (type PHRC¹²¹) « pour ne pas que ces tests très coûteux altèrent fortement l'enveloppe financière [du RIHN] ».

► Diffusion large de l'information auprès des professionnels et des patientes

La **LNCC**¹²² précise son intérêt de voir figurer à terme un complément méthodologique définissant un « *design* idéal » d'étude permettant d'obtenir des « conclusions fiables ». Elle questionne également les « effets potentiels » des conclusions du rapport sur une baisse d'intérêt et de financement venant de ses adhérents au profit de la médecine génomique. Dans ce contexte d'utilité clinique incertaine, la LNCC rappelle que « l'intérêt prioritaire à défendre est celui de l'information des patientes ».

Le **CNPP** précise également qu'il faudrait « largement diffuser et expliquer à la communauté médicale » les conclusions de la HAS, et en particulier auprès des pathologistes, qui seraient, selon eux, « la cible principale des demandes et des pressions exercées pour la réalisation des tests ».

Le **CNPP** appelle à définir en parallèle de cette évaluation les critères requis de sélection et de préparation de l'échantillon tumoral (analyse pré-analytique) pour réaliser ces tests génomiques dans de bonnes conditions¹²³.

¹²⁰ Notamment l'étude prospective OPTIGEN qui devait comparer les quatre SG n'a toujours pas démarré

¹²¹ Programme hospitalier de recherche clinique

¹²² Ligue nationale contre le cancer

¹²³ La HAS souhaite préciser que la réalisation de cette étape importante dans le processus est actuellement financée par un autre acte inscrit au RIHN (N005). Non spécifique du cancer du sein, il méritera d'être évalué dans sa globalité ultérieurement

4. Conclusions et perspectives

L'objectif de la présente évaluation est d'émettre un avis sur le bienfondé du remboursement pérenne par l'Assurance Maladie des deux SG de première génération (Oncotype Dx, MammaPrint) et des deux SG de deuxième génération (Prosigna, Endopredict) actuellement mises sur le marché, en vue d'optimiser la décision de CTA dans le cancer du sein de stade précoce. Ces tests bénéficient depuis avril 2016 d'un financement dérogatoire et transitoire avec une obligation de recueil de données cliniques dans le cadre du Référentiel des actes innovants hors nomenclatures (RIHN).

4.1 Conclusions

Dans le cadre de la présente évaluation, les cancers localement avancés, les tumeurs infracentimétriques hormonodépendantes RO+/HER2- sans invasion ganglionnaire (pN0) et les tumeurs non hormonodépendantes (RO-) ou HER2+ ou triple négatifs n'étaient pas des populations éligibles à une SG.

La HAS rappelle le **rôle essentiel des CCPT** pour la prise de décision de CTA **en contexte français** et considère que **les SG ne devraient pas être utilisées pour s'y substituer**.

Si selon certains experts du GT, une valeur pronostique (intrinsèque) des SG est possible, cette évaluation n'a pas permis de conclure en faveur d'une utilité clinique concernant ces quatre SG dans le cancer du sein de stade précoce. De plus, il a été retrouvé que la décision de CTA pouvait être divergente entre SG dans 20 à 25 % des cas pour un même échantillon de patientes. Par ailleurs, eu égard à leur objectif initial de développement et aux données analysées, les quatre SG ne peuvent revendiquer actuellement une valeur prédictive concernant l'efficacité d'un protocole de CTA.

Toutefois, cette évaluation a permis d'identifier une population de patientes pour qui la HAS considère, qu'en l'état actuel des connaissances, il existerait potentiellement et **dans le cadre de la recherche et de l'innovation**, un intérêt à évaluer l'utilisation des SG chez certaines patientes en **incertitude décisionnelle** de CTA au sein de la population préalablement ciblée par la HAS¹²⁴.

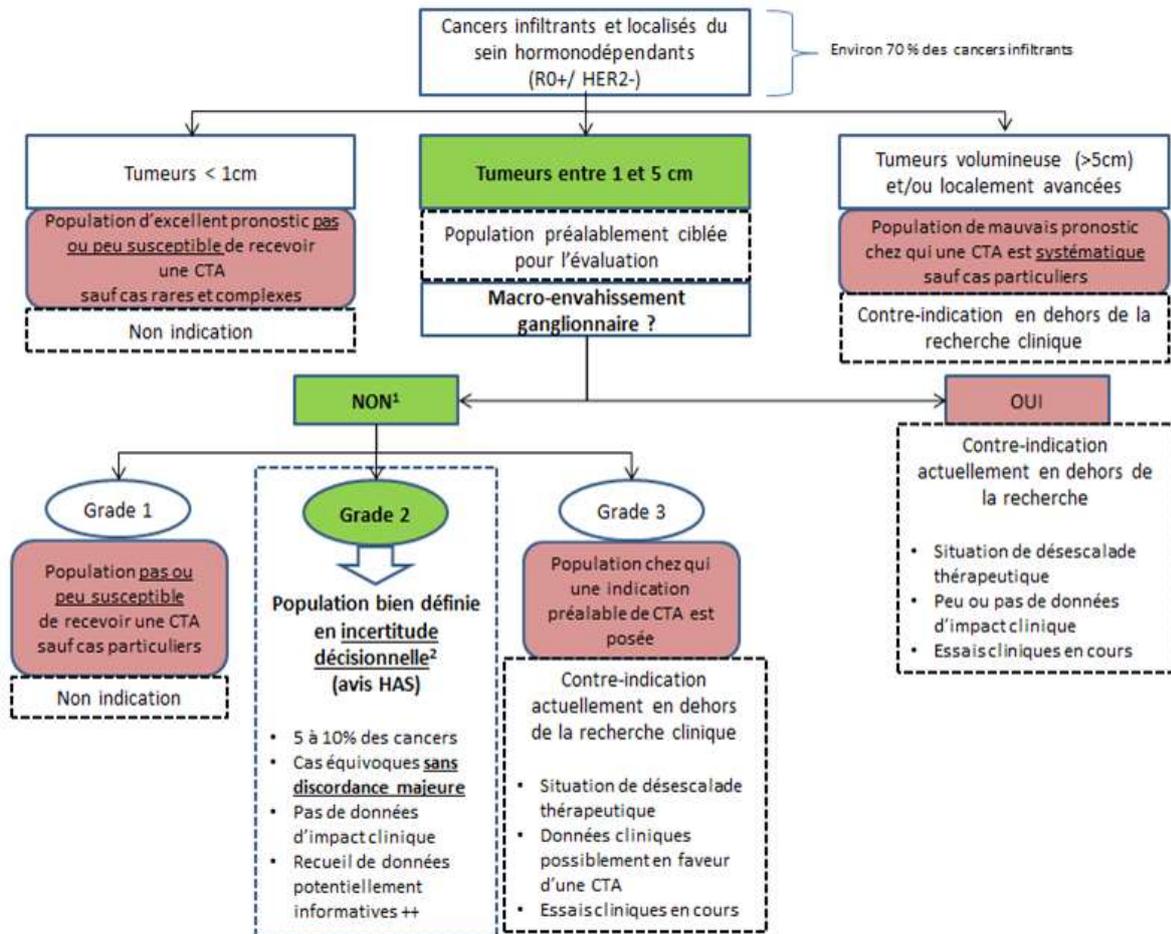
Ainsi, en conformité avec les recommandations de Saint Paul de Vence (2015), de Saint Gallen (2017) et la position du GT, la HAS propose d'encadrer cette utilisation au niveau d'une sous-population bien définie et plus restreinte de **cas équivoques** ayant un degré de prolifération **intermédiaire** (environ 5 à 10 % des cancers infiltrants). Ces patientes présentent un **adénocarcinome infiltrant** (préférentiellement canalaires), **pT1c-2 pN0-N1mi**¹²⁵ **de grade 2 et sans indication préalable de CTA au vu des autres critères de mauvais pronostic**¹²⁶. (cf. Figure 4).

¹²⁴ Cancer infiltrant du sein hormonodépendant RO+/HER2- de stade pT1c-pT2 pN0-1 (40 % des cancers infiltrants)

¹²⁵ Envahissement micrométastatique

¹²⁶ Âge jeune et présence d'emboles lymphovasculaires

Figure 4 : Etat des lieux et identification de la population en incertitude décisionnelle susceptible d'être explorée par une SG dans le cadre de la recherche (en vert)



1. En dehors des cas complexes d'incertitude liés à une discordance clinico-pathologique majeure (algorithme décisionnel validé préférable)
2. Vérifier l'absence d'indication préalable de CTA sur les autres critères de mauvais pronostic : âge jeune, embolies lymphovasculaires

La décision de prescription d'une SG pour les patientes éligibles devrait être **collégiale** (analyse des CCPT en réunion de concertation pluridisciplinaire) et après **accord de la patiente** (après une information claire, loyale et appropriée du professionnel et recueil de ses préférences en matière de choix thérapeutique).

La HAS considère que dans l'attente des résultats des trois essais en cours de réalisation¹²⁷ dans des populations à **plus haut risque clinique** (notamment N+), leur utilisation **ne peut être actuellement indiquée** dans l'optique d'une **désescalade thérapeutique** chez les patientes :

- qui présentent un envahissement ganglionnaire macrométastatique (patientes N+) ;
- qui présentent une indication préalable de CTA sur des critères de mauvais pronostic.

La HAS considère par ailleurs que leur utilisation **n'est actuellement pas indiquée** pour les cas complexes et hétérogènes d'incertitude décisionnelle (**discordance clinico-histo-pathologique majeure**). Le recours préférentiel à un **modèle de prédiction** du risque individuel qui serait à valider en contexte français serait un préalable indispensable. Ce type d'algorithme décisionnel pourrait permettre d'estimer la valeur ajoutée d'un score génomique dans l'optique de l'intégrer aux CCPT si son intérêt était démontrée.

¹²⁷ RXPONDER, WSG ADAPT (RO+/HER2-) et OPTIMA

La HAS considère qu'il est nécessaire de **rassurer le public et les patientes** sur la problématique française de l'administration de CTA en cas de cancer de stade précoce RO+/HER2- qui n'est pas celle du surtraitement de patientes qui seraient normalement peu susceptibles d'en recevoir.

La HAS considère que ses préconisations devraient être toutefois largement diffusées auprès des professionnels concernés afin de limiter le **mésusage** possible des SG :

- risque de sous-traitement par désescalade thérapeutique chez des patientes ayant une indication préalable de CTA ;
- risque d'escalade thérapeutique non pertinente chez les patientes normalement peu susceptibles de recevoir une CTA.

4.2 Perspectives

La HAS considère que le recueil des données dans le cadre du RIHN n'a pas contribué à l'évaluation des SG¹²⁸, mais plutôt, et dans la limite de leur exhaustivité, à observer leurs conditions d'utilisation en vie réelle par SG. La HAS préconise **une amélioration de la qualité du recueil de données cliniques relatives à leur utilisation qui devra être exhaustif et scientifiquement contributif en vue d'une prochaine évaluation technologique.**

A cet effet, le **schéma d'étude** nécessaire pour une démonstration formelle de l'utilité clinique d'une SG requerrait un essai prospectif randomisé à quatre bras (double randomisation).

A défaut, la HAS propose un schéma d'étude alternatif, **adaptatif** en deux temps :

- **Phase préliminaire observationnelle** (phase 1) estimant la valeur ajoutée « **analytique** »¹²⁹ des valeurs de score génomiques obtenues **en insu des CCPT** pour chaque signature dans cette population étroite et bien définie ;
- **Phase interventionnelle** (phase 2) : si la valeur ajoutée analytique est suffisante, il serait nécessaire d'attribuer alors **exclusivement** la CTA sur la base du score génomique obtenu **en insu des CCPT** par rapport au seuil de décision thérapeutique¹³⁰ du test. L'estimation de l'intérêt thérapeutique de la CTA guidée par une SG dans cette population serait établie au regard de l'importance clinique de la différence de **survie sans maladie invasive à cinq ans** entre les deux groupes.

Au-delà de ce travail, la HAS considère comme indispensable que soit élaboré, de préférence par l'INCa et en collaboration avec les professionnels concernés, **un document commun de recommandations nationales de bonne pratique sur les indications de CTA dans le cancer du sein RO+/HER2- de stade précoce.** Ce document viserait à réduire une part de l'hétérogénéité des prescriptions actuelles de CTA entre les centres d'oncologie. Ce travail d'harmonisation concernant les indications et les non indications établies de CTA permettrait de poursuivre également l'identification des situations d'incertitude décisionnelle pour les cliniciens (zone grise).

Pour résoudre les cas complexes d'incertitude décisionnelle, la HAS considère qu'un algorithme décisionnel d'aide à la décision intégrant les nombreux CCPT disponibles dans un modèle est devenu indispensable. Certains modèles existent et sont bien avancés¹³¹. Ceux-ci pourraient être adaptés puis validés en contexte français ou bien être développés *de novo via* des bases de données françaises récentes et de bonne qualité.

L'estimation du **risque de récurrence** des patientes en incertitude décisionnelle **sous HTA exclusive** pourrait être explorée à l'aide **d'études pronostiques** sur des données françaises récentes et de bonne qualité.

¹²⁸ L'étude comparative OPTIGEN n'a pas démarré à ce jour

¹²⁹ Étendue de la dispersion/distribution des valeurs de score génomique entre patientes (sans référence à aucun seuil de décision thérapeutique)

¹³⁰ Si valeur sous le seuil → HTA exclusive ; si valeur au-dessus du seuil → CTA + HTA

¹³¹ Par exemple, le modèle de prédiction PREDICT

Trois essais en cours de réalisation¹³² sont attendus dans des populations ciblées pour l'évaluation chez qui il existe une indication préalable de CTA (notamment les patientes N+). Il n'existe actuellement pas d'étude d'utilité clinique en cours pour les tests Endopredict et Mammaprint.

¹³² **Test Oncotype Dx** : essais RxPONDER et WSG ADAPT, **test Prosigna** : essai OPTIMA

Références bibliographiques

1. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69(3):89-95.
2. Bossuyt PM, Parvin T. Evaluating biomarkers for guiding treatment decisions. *EJIFCC* 2015;26(1):63-70.
3. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLoS Med* 2012;9(5):e1001216.
4. Haute Autorité de Santé. Test compagnon associé à une thérapie ciblée : définitions et méthode d'évaluation. Guide méthodologique. Saint-Denis La Plaine: HAS; 2014. https://www.has-sante.fr/portail/upload/docs/application/pdf/2014-04/guide_meth_court_test_cpagnon_vd.pdf
5. Ballman KV. Biomarker: Predictive or Prognostic? *Journal of Clinical Oncology* 2015;33(33):3968-71.
6. Pletcher MJ, Pignone M. Evaluating the clinical utility of a biomarker: a review of methods for estimating health impact. *Circulation* 2011;123(10):1116-24.
7. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clin Chem* 2012;58(12):1636-43.
8. Moons KG. Criteria for scientific evaluation of novel markers: a perspective. *Clin Chem* 2010;56(4):537-41.
9. Michiels S, Ternes N, Rotolo F. Statistical controversies in clinical research: prognostic gene signatures are not (yet) useful in clinical practice. *Ann Oncol* 2016;27(12):2160-7.
10. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, *et al.* The evaluation of genomic applications in practice and prevention (EGAPP) initiative: methods of the EGAPP Working Group. *Genet Med* 2009;11(1):3-14.
11. Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, *et al.* GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *Journal of clinical epidemiology* 2016;76:89-98.
12. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;350:g7647.
13. Croft P, Altman DG, Deeks JJ, Dunn KM, Hay AD, Hemingway H, *et al.* The science of clinical practice: disease diagnosis or patient prognosis? Evidence about "what is likely to happen" should shape clinical practice. *BMC Med* 2015;13:20.
14. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol* 2009;62(4):364-73.
15. Mustafa RA, Wiercioch W, Cheung A, Prediger B, Brozek J, Bossuyt P, *et al.* Decision making about healthcare-related tests and diagnostic test strategies. Paper 2: a review of methodological and practical challenges. *J Clin Epidemiol* 2017;92:18-28.
16. Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. *Ann Intern Med* 2011;154(4):253-9.
17. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, *et al.* Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013;346:e5793.
18. Institute of Medicine. Evolution of translational Omics: lessons learned and the path forward. Washington: IOM; 2012.
19. Baker SG, Kramer BS. Evaluating surrogate endpoints, prognostic markers, and predictive markers: Some simple themes. *Clin Trials* 2015;12(4):299-308.
20. Moons KG, de Groot JA, Linnet K, Reitsma JB, Bossuyt PM. Quantifying the added value of a diagnostic test or marker. *Clin Chem* 2012;58(10):1408-17.
21. Gleiss A, Zeillinger R, Braicu EI, Trillsch F, Vergote I, Schemper M. Statistical controversies in clinical research: the importance of importance. *Ann Oncol* 2016;27(7):1185-9.
22. American Society of Clinical Oncology, Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, *et al.* Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology clinical practice guideline. *J Clin Oncol* 2016;34(10):1134-50.
23. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, *et al.* Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012;98(9):691-8.
24. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, *et al.* Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10(2):e1001381.
25. Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, *et al.* The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol* 2017;87:4-13.
26. Ferrante di Ruffano L, Dinnes J, Taylor-Phillips S, Davenport C, Hyde C, Deeks JJ. Research waste in diagnostic trials: a methods review evaluating the reporting of test-treatment interventions. *BMC medical research methodology* 2017;17(1):32.
27. Santé publique France. Projection de l'incidence et de la mortalité par cancer en France métropolitaine en 2017. Rapport technique. Saint-Maurice: SFP; 2017.
28. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68(1):7-30.

29. Mamounas EP, Tang G, Liu Q. The importance of systemic therapy in minimizing local recurrence after breast-conserving surgery: the NSABP experience. *J Surg Oncol* 2014;110(1):45-50.
30. Tao JJ, Visvanathan K, Wolff AC. Long term side effects of adjuvant chemotherapy in patients with early breast cancer. *Breast* 2015;24 Suppl 2:S149-53.
31. Early Breast Cancer Trialists' Collaborative Group, Peto R, Davies C, Godwin J, Gray R, Pan HC, *et al.* Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet* 2012;379(9814):432-44.
32. Institut de recherche et documentation en économie de la santé. L'accès aux soins en cancérologie : évolution de l'offre et recours aux soins entre 2005 et 2012. *Questions d'économie de la santé* 2017(221):1-8.
33. Fekih M, Petit T, Zarca D, Guinebretière JM, André F, Pierga JY, *et al.* Utilisation de référentiels et hétérogénéité décisionnelle des indications de chimiothérapie adjuvante dans les cancers du sein exprimant les récepteurs hormonaux, HER2-négatifs: résultats d'un sondage national en France. *Bull Cancer* 2014;101(10):918-24.
34. Institut national du cancer, Société française de sénologie et de pathologie mammaire. uPA/PAI-1, Oncotype DX™, MammaPrint®. Valeurs pronostique et prédictive pour une utilité clinique dans la prise en charge du cancer du sein. Boulogne-Billancourt: INCA; 2013. <http://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Rapport-integral-UPA-PAI-1-ONCOTYPE-DXtm-MAMMAPRINT-R-Valeurs-pronostique-et-predictive-pour-une-utilite-clinique-dans-la-prise-en-charge-du-cancer-du-sein>
35. Institut national d'excellence en santé et en services sociaux. Utilisation d'EndoPredictMC et de ProsignaMC dans les cas de cancer du sein invasif précoce : INESSS; 2018. https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Oncologie/INESSS_EndoPredict-et-Prosigna_vs_Oncotype.pdf
36. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415(6871):530-6.
37. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351(27):2817-26.
38. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;27(8):1160-7.
39. Filipits M, Rudas M, Jakesz R, Dubsy P, Fitzal F, Singer CF, *et al.* A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res* 2011;17(18):6012-20.
40. Trosman JR, Van Bebber SL, Phillips KA. Coverage policy development for personalized medicine: private payer perspectives on developing policy for the 21-gene assay. *Journal of oncology practice* 2010;6(5):238-42.
41. National Comprehensive Cancer Network. Breast cancer. Version 2.2018. Fort Washington: NCCN; 2018.
42. Curigliano G, Burstein HJ, E PW, Gnant M, Dubsy P, Loibl S, *et al.* De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol* 2017;28(8):1700-12.
43. European Group on Tumor Markers, Duffy MJ, Harbeck N, Nap M, Molina R, Nicolini A, *et al.* Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). *Eur J Cancer* 2017;75:284-98.
44. American Society of Clinical Oncology, Krop I, Ismaila N, Andre F, Bast RC, Barlow W, *et al.* Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage-invasive breast cancer: American Society of Clinical Oncology clinical practice guideline focused update. *J Clin Oncol* 2017;35(24):2838-47.
45. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, *et al.* Tailoring therapies: improving the management of early breast cancer: Saint Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015;26(8):1533-46.
46. Sixièmes recommandations pour la pratique clinique de la prise en charge des cancers du sein de Nice - Saint Paul de Vence [En ligne] 2015. <http://www.cours-rpc-nice-saintpaul.fr/wp-content/uploads/2017/01/RPC-2015-RESUME-SITE.pdf>
47. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Thresholds for therapies: highlights of the Saint Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. *Ann Oncol* 2009;20(8):1319-29.
48. European Society for Medical Oncology, Aebi S, Davidson T, Gruber G, Cardoso F. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2011;22(Suppl 6):vi12-24.
49. National Comprehensive Cancer Network. Breast cancer. Version 1.2018. Fort Washington: NCCN; 2018.
50. Bachelot T, Bouzid K, Delozier T, Lortholary A, Petit T. État des lieux et mise à jour des traitements systémiques adjuvants. *Oncologie* 2011;13:698-702.
51. Spielmann M. Perception du pronostic des tumeurs mammaires pT1a,b pN0 par la communauté oncologique française : résultats de l'enquête nationale EURISTIC. *Bulletin du cancer* 2016;103(2):154-63.
52. Dalenc F, Penault-Llorca F, Cohen M, Houvenaeghel G, Piat JM, Liegeois P, *et al.* Daily practice management of pT1a-b pN0 breast carcinoma: a prospective french ODISSEE cohort study. *Clin Breast Cancer* 2017;17(2):107-16.
53. Vaz-Luis I, Ottesen RA, Hughes ME, Mamet R, Burstein HJ, Edge SB, *et al.* Outcomes by tumor subtype and treatment pattern in women with small, node-negative

- breast cancer: a multi-institutional study. *J Clin Oncol* 2014;32(20):2142-50.
54. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Ann Oncol* 2007;18(7):1133-44.
55. European Society for Medical Oncology, Kataja V, Castiglione M. Primary breast cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol* 2009;20 Suppl 4:10-4.
56. European School of Oncology and Breast, European Society for Medical Oncology, Cardoso F, Costa A, Norton L, Senkus E, *et al.* ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). *Ann Oncol* 2014;25(10):1871-88.
57. Chereau E, Coutant C, Gligorov J, Lesieur B, Antoine M, Darai E, *et al.* Discordance with local guidelines for adjuvant chemotherapy in breast cancer: reasons and effect on survival. *Clin Breast Cancer* 2011;11(1):46-51.
58. Sabiani L, Houvenaeghel G, Heinemann M, Reyat F, Classe JM, Cohen M, *et al.* Breast cancer in young women: Pathologic features and molecular phenotype. *Breast* 2016;29:109-16.
59. Blum JL, Flynn PJ, Yothers G, Asmar L, Geyer CE, Jr., Jacobs SA, *et al.* Anthracyclines in early breast cancer: the ABC trials-USOR 06-090, NSABP B-46-I/USOR 07132, and NSABP B-49 (NRG Oncology). *J Clin Oncol* 2017;35(23):2647-55.
60. Haute Autorité de Santé, Institut national du cancer. Tumeur maligne, affection maligne du tissu lymphatique ou hématopoïétique. Cancer du sein. Guide ALD. Saint-Denis La Plaine: HAS; 2010.
https://www.has-sante.fr/portail/upload/docs/application/pdf/2010-02/ald_30_gm_ksein_vd.pdf
61. 5es Recommandations francophones pour la pratique clinique (RPC) « cancers du sein » / Nice – Saint-Paul-de-Vence 2013. *Oncologie* 2013;15(12).
62. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, *et al.* Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst* 2011;103(22):1656-64.
63. Luporsi E, Andre F, Spyrtos F, Martin PM, Jacquemier J, Penault-Llorca F, *et al.* Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat* 2012;132(3):895-915.
64. Denkert C, Budczies J, von Minckwitz G, Wienert S, Loibl S, Klauschen F. Strategies for developing Ki67 as a useful biomarker in breast cancer. *Breast* 2015;24(Suppl 2):S67-72.
65. Penault-Llorca F, Radosevic-Robin N. Ki67 assessment in breast cancer: an update. *Pathology* 2017;49(2):166-71.
66. Filleron T, Chaltiel L, Jouve E, Cabarrou B, Gilhodes J, Lusque A, *et al.* Les nomogrammes dans la pratique clinique : méthodologie, limites et intérêts. *Bull Cancer* 2018;105(1):15-24.
67. Strasser-Weippl K, Goss PE. Competing risks in low-risk breast cancer. *Am Soc Clin Oncol Educ Book* 2013:32-9.
68. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, *et al.* The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* 2017;67(2):93-9.
69. Institut national du cancer. Cancer du sein infiltrant non métastatique. Questions d'actualité. Boulogne Billancourt: INCA; 2012.
http://www.e-cancer.fr/content/download/95914/1021156/file/RECOSEI_NINFILT12.PDF
70. Vickers AJ, Kattan MW, Daniel S. Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials* 2007;8:14.
71. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162(1):W1-73.
72. Moosdorff M, van Roozendaal LM, Strobbe LJ, Aebi S, Cameron DA, Dixon JM, *et al.* Maastricht Delphi consensus on event definitions for classification of recurrence in breast cancer research. *J Natl Cancer Inst* 2014;106(12).
73. Ong WL, Schouwenburg MG, van Bommel ACM, Stowell C, Allison KH, Benn KE, *et al.* A standard set of value-based patient-centered outcomes for breast cancer: the International Consortium for Health Outcomes Measurement (ICHOM) initiative. *JAMA Oncol* 2017;3(5):677-85.
74. Institut Curie. Rapport de l'évaluation prévue au RIHN. Signatures d'expression génique dans le cancer du sein. N°537. Impact décisionnel et budgétaire du test Prosigna® dans le cancer du sein précoce. Paris: Institut Curie.
75. Agency for Healthcare Research and Quality. Technology assessment of molecular pathology testing for the estimation of prognosis for common cancers. Rockville: AHRQ; 2014.
https://www.ncbi.nlm.nih.gov/books/NBK285003/pdf/Books_helf_NBK285003.pdf
76. Canadian Agency for Drugs and Technologies in Health. Gene expression tests for women with early stage breast cancer: A review of clinical utility and cost-effectiveness. Ottawa: CADTH; 2017.
https://www.cadth.ca/sites/default/files/pdf/htis/2017/RC09_34%20Mic%20Tests%20for%20Breast%20Cancer%20Final.pdf
77. Agence canadienne des médicaments et des technologies de la santé. Examen de l'utilisation du test Oncotype DX pour le cancer du sein au stade précoce sans envahissement ganglionnaire. Ottawa: ACMTS; 2014.
https://cadth.ca/sites/default/files/pdf/htis/apr-2014/RC0524_RR_RiB_OncotypeDX_LNneg_f.pdf

78. Agence canadienne des médicaments et des technologies de la santé. Examen d'Oncotype DX dans le cancer du sein en stade précoce à envahissement ganglionnaire. Ottawa: ACMTS; 2014.
79. Australian Government Department of Health. Application n°1342.4. Gene expression profiling of 21 genes in breast cancer to quantify the risk of disease recurrence and predict adjuvant chemotherapy benefit. Canberra: MSAC; 2017.
80. Australian Government Department of Health. Application n°1473. 50 gene signature assay for predicting breast cancer recurrence. Canberra: MSAC; 2017.
81. European Network for Health Technology Assessment. MammaPrint®. Added value of using the gene expression signature test MammaPrint® for adjuvant chemotherapy decision-making in early breast cancer : EUnetHTA; 2018.
82. Centre fédéral d'expertise des soins de santé. MammaPrint® test for personalised management of adjuvant chemotherapy decisions in early breast cancer. KCE Reports 298. Brussels: KCE; 2018.
https://kce.fgov.be/sites/default/files/atoms/files/KCE_298_MammaPrint_tests_Report.pdf
83. Institute for Quality and Efficiency in Health Care. Biomarker-based tests to support the decision for or against adjuvant systemic chemotherapy in primary breast cancer. Cologne: IQWiG; 2016.
84. Lange S, Scheibler F, Fleer D, Windeler J. Interpretation of the results of the MINDACT study and consequent recommendations in the updated ASCO clinical practice guideline. *J Clin Oncol* 2018;36(4):429-30.
85. Harbeck N, Sotlar K, Wuerstlein R, Doisneau-Sixou S. Molecular and protein markers for clinical decision making in breast cancer: today and tomorrow. *Cancer Treat Rev* 2014;40(3):434-44.
86. Haute Autorité de Santé. Description générale de la procédure d'évaluation d'actes professionnels. Saint-Denis La Plaine: HAS; 2018.
https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-03/has_methode_generale_actes_08_03_2018.pdf
87. Paynter NP, Cook NR. Adding tests to risk based guidelines: evaluating improvements in prediction for an intermediate risk group. *BMJ* 2016;354:i4450.
88. American Heart Association, Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, *et al.* Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation* 2009;119(17):2408-16.
89. De Bin R, Herold T, Boulesteix AL. Added predictive value of omics data: specific issues related to validation illustrated by two case studies. *BMC Med Res Methodol* 2014;14:117.
90. Tajik P, Zafarmand MH, Zwinderman AH, Mol BW, Bossuyt PM. Development and evaluating multimarker models for guiding treatment decisions. *BMC medical informatics and decision making* 2018;18(1):52.
91. Janes H, Brown MD, Cramer MR, Miller DP, Barlow WE. Adjusting for covariates in evaluating markers for selecting treatment, with application to guiding chemotherapy for treating estrogen-receptor-positive, node-positive breast cancer. *Contemporary clinical trials* 2017;63:30-9.
92. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, *et al.* GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011;64(4):395-400.
93. European Network for Health Technology Assessment. Endpoints used for relative effectiveness assessment: clinical endpoints : EUnetHTA; 2015.
94. Hudis CA, Barlow WE, Costantino JP, Gray RJ, Pritchard KI, Chapman JA, *et al.* Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol* 2007;25(15):2127-32.
95. Gourgou-Bourgade S, Cameron D, Poortmans P, Asselain B, Azria D, Cardoso F, *et al.* Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials) dagger. *Ann Oncol* 2015;26(5):873-9.
96. Food and Drug Administration, U.S. Department of Health and Human Services. Guidance for industry clinical trial endpoints the approval of cancer drugs and biologics. Rockville: FDA; 2007.
<https://www.fda.gov/downloads/Drugs/Guidances/ucm071590.pdf>
97. European Medicines Agency. Guideline on the evaluation of anticancer medicinal products in man. London: EMA; 2013.
https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-4_en.pdf
98. European Society for Medical Oncology, Cherny NI, Sullivan R, Dafni U, Kerst JM, Sobrero A, *et al.* A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol* 2015;26(8):1547-73.
99. Sparano JA, Solin LJ. Defining the clinical utility of gene expression assays in breast cancer: the intersection of science and art in clinical decision making. *J Clin Oncol* 2010;28(10):1625-7.
100. European Medicines Agency. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man. London: EMA; 2012.
https://www.ema.europa.eu/documents/scientific-guideline/appendix-1-guideline-evaluation-anticancer-medicinal-products-man-methodological-consideration-using_en.pdf

101. Lacny S, Wilson T, Clement F, Roberts DJ, Faris P, Ghali WA, *et al.* Kaplan-Meier survival analysis overestimates cumulative incidence of health-related events in competing risk settings: a meta-analysis. *J Clin Epidemiol* 2018;93:25-35.
102. Penault-Llorca F, Coeffic D, Delozier T, Dohollou N, Freyer G, Gligorov J, *et al.* Le cancer du sein sans envahissement ganglionnaire. Au-delà des consensus internationaux : une approche pragmatique. *Bull Cancer* 2011;98(7):807-25.
103. Melvin JC, Purushotham AD, Garmo H, Pinder SE, Fentiman IS, Gillett C, *et al.* Progression of breast cancer following locoregional ipsilateral recurrence: importance of interval time. *Br J Cancer* 2016;114(1):88-95.
104. Zingarello A, Mazouni C, Rivera S, Mokdad-Adi M, Pistilli B. Prognostic assessment and systemic treatments of invasive local relapses of hormone receptor-positive breast cancer. *Breast* 2017;35:162-8.
105. European Society for Medical Oncology, Cardoso F, Harbeck N, Fallowfield L, Kyriakides S, Senkus E. Locally recurrent or metastatic breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2012;23 Suppl 7:vii11-9.
106. Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, *et al.* Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. *The Lancet Oncology* 2016;17(11):e510-e4.
107. Bonnetain F, Fiteni F, Efficace F, Anota A. Statistical challenges in the analysis of health-related quality of life in cancer clinical trials. *J Clin Oncol* 2016;34(16):1953-6.
108. Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *Lancet Oncol* 2015;16(1):e32-42.
109. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, *et al.* 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med* 2016;375(8):717-29.
110. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, *et al.* Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med* 2018.
111. Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W, *et al.* Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. *Breast Cancer Res Treat* 2017.
112. Bartlett JM, Bayani J, Marshall A, Dunn JA, Campbell A, Cunningham C, *et al.* Comparing breast cancer multiparameter tests in the OPTIMA prelim trial: no test is more equal than the others. *J Natl Cancer Inst* 2016;108(9).
113. Bueno-de-Mesquita JM, van Harten WH, Retel VP, van 't Veer LJ, van Dam FS, Karsenberg K, *et al.* Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). *The Lancet. Oncology* 2007;8(12):1079-87.
114. Drukker CA, Bueno-de-Mesquita JM, Retel VP, van Harten WH, van Tinteren H, Wesseling J, *et al.* A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *Int J Cancer* 2013;133(4):929-36.
115. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, *et al.* Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med* 2015;373(21):2005-14.
116. Stein RC, Dunn JA, Bartlett JM, Campbell AF, Marshall A, Hall P, *et al.* OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. *Health technology assessment* 2016;20(10):xxiii-xxix, 1-201.
117. Gluz O, Nitz UA, Christgen M, Kates RE, Shak S, Clemens M, *et al.* West german study group phase III plan B trial: first prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by central and local pathology assessment. *J Clin Oncol* 2016;34(20):2341-9.
118. Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* 1999;319(7223):1492-5.
119. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308(24):2594-604.
120. Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet* 2002;359(9318):1686-9.
121. Vervölgyi E, Kromp M, Skipka G, Bender R, Kaiser T. Reporting of loss to follow-up information in randomised controlled trials with time-to-event outcomes: a literature survey. *BMC medical research methodology* 2011;11:130.
122. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ (Clinical research ed.)* 2018;363:k4245.
123. Cherny NI, Dafni U, Bogaerts J, Latino NJ, Pentheroudakis G, Douillard JY, *et al.* ESMO-magnitude of clinical benefit scale version 1.1. *Ann Oncol* 2017;28(10):2340-66.
124. Maisonneuve H, Babany G. Données de vie réelle et recherche clinique : le complément des données des essais randomisés ? *Presse Med* 2015;44(6 Pt 1):586-9.
125. Sestak I, Buus R, Cuzick J, Dubsy P, Kronenwett R, Denkert C, *et al.* Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA oncology* 2018;4(4):545-53.
126. Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C, *et al.* Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the Arimidex, Tamoxifen, Alone or in Combination trial. *J Clin Oncol* 2008;26(7):1059-65.

Fiche descriptive

Intitulé	Descriptif
Méthode de travail	Évaluation d'une technologie de santé
Date de mise en ligne	Février 2019
Date d'édition	Uniquement disponible sous format électronique sur www.has-sante.fr
Objectif(s)	Évaluer l'utilité clinique des signatures génomiques pour décider d'une chimiothérapie adjuvante dans une population à risque intermédiaire de récurrence tumorale
Professionnel(s) concerné(s)	Cf. chapitre 2.6
Demandeur	Autosaisine de la HAS
Promoteur	Haute Autorité de santé (HAS), service évaluation des actes professionnels (SEAP)
Pilotage du projet	Chef de service : Cédric CARBONNEIL, adjoint au chef de service : Nadia SQUALLI Secrétariat : Lina BISCOSI, assistante, SEAP
Participants	Expertise externe à la HAS : Madame le Docteur Carine BELLERA, Monsieur le Docteur Jean Marie BOHER, Madame le Docteur Patricia DE CREMOUX, Monsieur le Docteur Frédéric FITENI, Madame le Docteur Sophie GOURGOU, Madame le Docteur Séverine GUIU, Monsieur le Docteur Jonathan LOPEZ, Madame le Docteur Elisabeth LUPORSI, Monsieur le Docteur Gaëtan MAC GROGAN, Monsieur le Dr Jean-Baptiste SOURAUD, Madame le Docteur Christine TUNON DE LARA, Madame Véronique WOLF (cf. chapitre 2.6)
Recherche documentaire	De janvier 2002 à août 2018 (stratégie de recherche documentaire décrite en annexe 13) Réalisée par Sophie DESPEROUX, documentaliste, avec l'aide de Yasmine LOMBRY, assistante documentaliste, sous la responsabilité de Frédérique PAGES, chef du service documentation - veille, et Christine DEVAUD, adjointe au chef de service
Auteurs de l'argumentaire	Yann CHAMBON, chef de projet, SEAP, sous la responsabilité de Nadia SQUALLI, adjoint au chef de service, SEAP
Validation	Examen par la Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (CNEDiMTS) : novembre 2018 Collège de la HAS : janvier 2019
Autres formats	Pas d'autre format que le format électronique disponible sur www.has-sante.fr
Documents d'accompagnement	Annexes au rapport d'évaluation, avis et décision HAS (janvier 2019) disponibles sur www.has-sante.fr



Toutes les publications de la HAS sont téléchargeables sur
www.has-sante.fr