



HAUTE AUTORITÉ DE SANTÉ

ANNEXES AU RAPPORT D'ÉVALUATION TECHNOLOGIQUE

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

Janvier 2019

Sommaire

| | |
|--|----|
| Annexe 1. Tableau synoptique des caractéristiques techniques des relatives aux SG..... | 4 |
| Annexe 2. Classification histopathologique pTNM (UICC, 8th édition, 2017)..... | 5 |
| Annexe 3. Caractéristiques des patientes testées selon les données transmises dans le cadre du RIHN (2016 – 2017) | 7 |
| Annexe 4. Etudes d'impact clinique en cours de réalisation | 8 |
| Annexe 5. Niveau de sévérité d'un évènement indésirable (EI) selon la classification CTCAE V5.0 (2017) . | 9 |
| Annexe 6. Tableaux de données de l'étude MINDACT {Cardoso, 2016 #161}..... | 10 |
| Annexe 7. Tableaux de données de l'étude TAILORx {Sparano, 2018 #159}..... | 13 |
| Annexe 8. Tableaux de données de l'étude ancillaire à PLAN B {Nitz, 2017 #22}..... | 17 |
| Annexe 9. Tableaux de données de l'étude OPTIMA Prelim {Bartlett, 2016 #177}..... | 19 |
| Annexe 10. Compte rendu de la réunion du GT du 02 mai 2018..... | 20 |
| Annexe 11. Compte rendu de la réunion du GT du 02 juillet 2018 | 27 |
| Annexe 12. Retour <i>in extenso</i> des parties prenantes | 35 |
| Annexe 13. Recherche documentaire | 62 |

Annexe 1. Tableau synoptique des caractéristiques techniques des relatives aux SG

| | Oncotype Dx (1^{ère} génération) | Mammaprint (1^{ère} génération) | Prosigna (2^{ème} génération) | Endopredict (2^{ème} génération) |
|--------------------------------|---|---|---|--|
| Fabricant | Genomic Health | Agendia | Nanostring | Myriad |
| Lieu du dosage | Plateforme centralisée (Etats-Unis) | Deux plateformes centralisées (Pays Bas, Etats-Unis) Une plateforme décentralisée en France (Institut Curie) | Multiplates plateformes décentralisées localement dans des centres français | Multiplates plateformes décentralisées localement dans des centres français |
| Nombre de gènes testés | 21 (5 de référence) | 70 | 50 (4 de référence) | 12 (4 de référence) |
| Matériel tumoral | Tissu fixé en paraffine | Tissu frais (ou plus récemment sur tissu fixé en paraffine) | Tissu fixé en paraffine | Tissu fixé en paraffine |
| Méthode d'analyse | qRT-PCR | Puce à ADN (ou plus récemment séquençage haut débit) | Hybridation ARNm nCounter® | qRT-PCR |
| Score génomique | <i>Recurrence score (RS)</i> Score de 0 à 100 | Haut ou bas risque génomique Score de -1 à +1 | <i>Risk of recurrence (ROR-PT)</i> Score de 0 à 100 | EPclin Score de 1 à 6 |
| CCPT | Non | Non | Oui Taille tumorale (T > ou < 2cm) Ganglions (N0 ou N1) | Oui Taille tumorale (pT1ab, pT1c, pT2, pT3) Ganglions (N0, N1, N2, N3) |
| Catégories de risque génomique | Faible/intermédiaire /élevé | faible/élevé | Faible/intermédiaire/ élevé | faible/élevé |
| Marquage CE | Non | Oui | Oui | Oui |

qRT-PCR : reverse transcriptase quantitative ;

Annexe 2. Classification histopathologique pTNM (UICC, 8th édition, 2017)

pT - Tumeur primitive

| | |
|-------------|--|
| pTx | Renseignements insuffisants |
| pT0 | Absence de tumeur primitive retrouvée |
| pTis | Carcinome <i>in situ</i> |
| pT1 | Tumeur de moins de 2 cm dans son plus grand axe |
| pT1a | Tumeur entre 1 et 5 mm |
| pT1b | Tumeur entre 5 mm et 1 cm |
| pT1c | Tumeur entre 1 cm et 2 cm |
| pT2 | Tumeur entre 2 et 5 cm dans son plus grand axe |
| pT3 | Tumeur de plus de 5 cm dans son plus grand axe. |
| pT4 | Tumeur de toute taille avec un envahissement de la cage thoracique et/ou de la peau (ulcération ou nodule) |
| pT4a | Tumeur envahissant la cage thoracique (à l'exclusion du muscle pectoral) |
| pT4b | Ulcération, nodules cutanés homolatéraux, œdème (incluant une peau d'orange) |
| pT4c | pT4a et pT4b inclus |
| pT4d | Carcinome inflammatoire |

pN - Adénopathies régionales et M- Métastases à distance

| | |
|--------------|---|
| pNx | Ne peut pas être évalué (curage ganglionnaire ancien ou non réalisé) |
| pN0 | Absence de métastase ganglionnaire lymphatique régionale |
| pN1 | Micrométastase, 1 à 3 ganglions axillaires homolatéraux métastatiques et/ou mammaires internes non cliniquement détectables (technique du ganglion sentinelle) |
| pN1mi | Micrométastases (entre 0,2 mm ou 200 cellules et 2 mm) |
| pN1a | 1 à 3 ganglions axillaires homolatéraux métastatiques (incluant au moins une macrométastase de plus de 2 mm) |
| pN1b | Envahissement des ganglions mammaires internes |
| pN1c | 1 à 3 ganglions axillaires homolatéraux métastatiques (incluant au moins une macrométastase de plus de 2 mm) et envahissement des ganglions mammaires internes |
| pN2 | 4 à 9 ganglions axillaires homolatéraux métastatiques ou adénopathies métastatiques mammaires internes cliniquement détectables en absence de métastases axillaires |
| pN2a | 4 à 9 ganglions axillaires homolatéraux métastatiques (incluant au moins une macrométastase de plus de 2 mm) |
| pN2b | Adénopathies métastatiques mammaires internes cliniquement détectables en absence de métastases axillaires |
| pN3 | |
| pN3a | Au moins 10 ganglions axillaires homolatéraux métastatiques (incluant au moins une macrométastase de plus de 2 mm) ou métastases infra-claviculaires |
| pN3b | Adénopathies métastatiques mammaires internes homolatérales cliniquement détectables en présence de métastases axillaires associées ou au moins 3 ganglions axillaires homolatéraux |

| | |
|-------------|--|
| pTx | Renseignements insuffisants |
| | métastatiques et envahissement des ganglions mammaires internes non cliniquement détectables mais micro ou macrométastases avec la technique du ganglion sentinelle |
| pN3c | Envahissement des ganglions supra-claviculaire homolatéraux |
| M0 | Pas de métastase à distance |
| M1 | Présence de métastase(s) à distance |

Annexe 3. Caractéristiques des patientes testées selon les données transmises dans le cadre du RIHN (2016 – 2017)

| | Oncotype Dx | Prosigna | Endopredict |
|---|--|--|---|
| Période de recrutement (durée du recueil RIHN) | Octobre 2015 à juillet 2016 (4 mois) | Avril 2016 à août 2017 (16 mois) | Mai 2016 à décembre 2017 (19 mois) |
| Nombre de tests réalisés (centres recruteurs) | 882 (53 centres) | 571 (non précisé) | 302 en 2016 1 246 en 2017 (non précisé) |
| Nombre de plateformes de biologie concernées | 1 (plateforme centralisée aux USA) | 1 (plateforme de l'Institut Curie exclusivement) | 6 en 2016 9 en 2017 |
| Données manquantes | Non précisées | 60 % (patientes prises en charge hors Institut Curie) | 10 % en 2016 1 % en 2017 |
| Référentiel appliqué pour la prescription du test | pTx pN1-3 (selon les critères d'éligibilité du fabricant) | REMAGUS-UNICANCER ¹ | GFCO ² |
| Proportion de prescriptions hors référentiel local | 0,4 % homme 2 % HER2+ | 26 % | Quelques pT1a et pT3 |
| Proportion de prescription hors population ciblée (HAS) | Au moins 12-14 % (pT1ab, pT3, HER2+) | Non estimable | Au moins 12 % en 2016 (pT1ab, pT3) Au moins 14% en 2017 (pT1b, pT3, pN2) |
| Ganglions envahis pN1 | pN+ 29 % dont 21 % macroN+ | 12 % | 23 % en 2016 25 % en 2017 (dont 20 % macro N+) |
| Grade 2 histopronostique | 68 % | 81% | 82% en 2016 75% en 2017 |
| Grade 3 histopronostique | 18 % | 14 % | 6% en 2016 12 % en 2017 |
| Tumeurs non canalaies | 21% | Non précisé | 14 % en 2016 21 % en 2017 |
| Score génomique intermédiaire | 36% (RS 18 – 30) | 37 % | Non applicable |

¹ Référentiel de l'Institut Curie et de l'Institut Gustave Roussy

² Groupe francophone de cytogénomique oncologique

Annexe 4. Etudes d'impact clinique en cours de réalisation

| Mammaprint (Agendia) | Oncotype Dx (Genomic Health) | Endopredict (Myriad) | Prosigna (Nanostring) |
|--------------------------------|--|--------------------------------|---|
| AUCUN ESSAI PREVU | <p>RXPONDER (2011-X) (SWOG S1007) Soutenu par SWOG/NCI NCT01272037</p> <p><i>Utilité clinique du test sur la valeur prédictive d'efficacité en vue d'une désescalade de CTA.</i></p> <p><i>Tumeurs RO+/HER2- de stade précoce ou localement avancé (pT1-pT4 sauf inflammatoire), pN1a.</i></p> <p><i>Inclusions : 10 000</i></p> <p><i>Analyse finale prévue en 2016 puis reportée à 2022</i></p> | AUCUN ESSAI PREVU | <p>OPTIMA (2012-X) soutenu par le NHS ISRCTN42400492</p> <p><i>Utilité clinique du test en cas d'indication préalable de CTA en vue d'une désescalade thérapeutique</i></p> <p><i>Tumeurs IRO+/HER2- de stade précoce pN0-1 et localement avancé pN2.</i></p> <p><i>Inclusions : 4 500</i></p> <p><i>Analyse finale prévue au-delà de 2020</i></p> |
| | <p>WSG ADAPT (RO+/HER2-) (2012-X) soutenu par Genomic Health et AOK Rheinland/Hamburg NCT01779206</p> <p><i>Utilité clinique d'Oncotype Dx en association avec le Ki67 en vue d'une désescalade de la CTA et étude de la valeur prédictive d'hormonosensibilité.</i></p> <p><i>Tumeurs RO+/HER2- de stade précoce et localement avancé, pN0-N1.</i></p> <p><i>Analyse finale prévue en 2019/2020.</i></p> <p><i>Inclusions prévues : 4 000</i></p> | | |

Annexe 5. Niveau de sévérité d'un évènement indésirable (EI) selon la classification CTCAE V5.0 (2017)

| | GRADE 1 Léger | GRADE 2 Modéré | GRADE 3 Sévère | GRADE 4 Pronostic vital en jeu | GRADE 5 Décès |
|--|--|---|--|---|-------------------------------------|
| Sévérité de l'évènement indésirable | Asymptomatique ou symptômes légers, nécessite une simple observation clinique ou une prise en charge diagnostique mais sans intervention correctrice | Intervention correctrice minime et locale ou non invasive. Limitation dans les activités quotidiennes du patient | Hospitalisation ou prolongation de l'hospitalisation en rapport avec l'évènement indésirable Limitation dans l'autonomie essentielle du patient | Nécessite une intervention correctrice en urgence | Décès lié à l'évènement indésirable |

Annexe 6. Tableaux de données de l'étude MINDACT {Cardoso, 2016 #161}

| | |
|--|---|
| Sponsors de l'étude | Étude européenne de financement public (Commission européenne et organisation EORTC). Tests fournis par le fabricant AGENDIA. |
| Déclarations des liens d'intérêt | Sur 34 signataires, cinq co-auteurs ont déclarés une situation à risque majeure de conflits d'intérêts dont deux salariés du fabricant et un détenteur du brevet du test Mammaprint. |
| Informations générales | <p>Période de recrutement : 2007 – 2011, médiane de suivi de cinq ans (gel de la base en mars 2016).</p> <p>Inclusion de 59 % des 11 288 patientes éligibles. Non inclusion pour échecs techniques (17,3 %) ou refus (8 %) :</p> <ul style="list-style-type: none"> deux registres de suivi (si risques concordants) entre le risque clinique C et génomique G (C haut/G haut sous CTA/HT ou C bas/G bas sous CTA) ; double randomisation des traitements CT/no CT (si risques discordants) : C haut/ G-bas et C bas/G haut. <p>Le risque clinique estimé par l'algorithme Adjuvant ! (version 8 modifiée avec HER2).</p> <p>Si C haut = indication théorique de CTA pour les auteurs.</p> |
| Caractéristiques de la population incluse | <p>N= 6 693 patientes incluses.</p> <p>Patientes 18 - 70 ans ; pT1-2³, pN0-1; <u>tous profils moléculaires confondus</u> : 9,6 % triple négatif et 9,5 % HER2 +.</p> <p>Population d'intérêt : 81 % tumeurs lumineales (n= 5 408) :</p> <ul style="list-style-type: none"> pN0 (78 %) (n = 4 225) ; pN1 (22 %) dont pN1mi. <p>Deux groupes à risque concordants pour assignation ;</p> <ul style="list-style-type: none"> C bas /G bas : assignation HT exclusive (n= 2 634 en statut quo) ; C haut / G haut : assignation HT/CT (n= 1 873 en statut quo). <p>Deux groupes à risque discordants pour double randomisation (n= 2 187) :</p> <ul style="list-style-type: none"> C haut / G bas (n= 1 497) = population d'analyse primaire pour désescalade ; C bas / G haut (n=690). <p>Caractéristiques clinico-pathologiques des patientes (après correction des erreurs de risque / per corrected) :</p> <ul style="list-style-type: none"> corrélation forte entre le grade 3 et le risque génomique G : <ul style="list-style-type: none"> 11 fois plus de grade 3 entre G haut et G bas chez les patientes C bas ; 2,5 fois plus de grade 3 entre G haut et G bas chez les patientes C haut ; 75 % de grade 3 chez les risques concordants C haut / G haut. <p>Corrélation forte entre le risque génomique G et le profil moléculaire non luminal (11 à 24 fois plus de G haut en cas de C haut à tumeur non lumineale.</p> |
| Populations | Dans les deux groupes à risque discordants (et randomisés) : |

³ Taux négligeable de pT3 = 1,2%

| | |
|---|---|
| <p>d'analyse et qualité de conduite de l'étude</p> | <p>Essai en ouvert du groupe d'allocation et du score génomique. Traitements laissés à l'appréciation des investigateurs.</p> <p>Taux d'écarts majeurs au protocole⁴ : 24,4 % (534 / 2 187) :</p> <ul style="list-style-type: none"> • C bas / G haut = 35 % ; • C haut / G bas = 20 %. <p>Taux de sorties prématurées (censures) : non rapporté par les auteurs mais calculable...</p> <p>Informations sur la nature des traitements reçus (HT et CT) et les arrêts prématurés de traitement : non rapportées.</p> <p>L'analyse en intention de traiter (ITT) est à haut risque de biais (devant le taux important d'écarts majeurs au protocole pour une recherche de non infériorité)</p> <p>L'analyse en per protocole (PP) et son analyse de sensibilité (PPS)⁵ prévues dans le plan d'analyse statistique si nombreuses violations au protocole étaient préférables.</p> <p>Niveau d'adhésion à la décision génomique (Figure 1 et tableau S11) :</p> <ul style="list-style-type: none"> • C haut/G bas : 11 % refus de la désescalade (85 / 748) partagé entre le refus de la patiente et du clinicien. <p>C bas/G haut : 22 % refus de la CT (76 / 344) souvent par la patiente.</p> |
| <p>Résultats</p> | <p>Question 1 : Utilisation de Mammaprint en substitution des CCP habituels (comparaison du devenir des cas à risque discordant).</p> <p>Comparaison entre la <u>stratégie</u> basée sur la génomique⁶ (G risk) et celle basée sur la clinique⁷ (C risk) : optimisation de la décision :</p> <ul style="list-style-type: none"> • résultat contenu p. 724 de la publication et dans la figure S1 des suppléments ; • une seule analyse de ce type mais peu appropriée et ne concernant la population d'intérêt de l'évaluation : <ul style="list-style-type: none"> ▸ tous profils moléculaires et ganglionnaires confondus ; ▸ en ITT (et non en PP qui est à préférer à cause des écarts au protocole) ; ▸ en incluant également les patientes à risque concordant en statut quo décisionnel (ce qui dilue les différences) et pousse les auteurs à devoir pondérer artificiellement certaines estimations par un facteur 2... ; ▸ la comparaison sur la survie sans maladie (SSM) plus pertinente en contexte adjuvant, est manquante ; ▸ stratégie clinique 95 % vs stratégie génomique 94,7 % sur la survie sans métastase à distance (SSMD) à cinq ans ; ▸ pas de preuve de l'optimisation décisionnelle avec la génomique (bénéfice observé de 0,3 % de SSMD à cinq ans avec Adjuvant !). Aucune comparaison rapportée limitée aux patientes à risque discordant dans la population d'intérêt de l'évaluation ; • estimation par la HAS de la différence absolue Δ entre les deux stratégies (calcul limité aux patientes à risque discordant) = Δ ponctuel entre - 0,8 % et - 1,5 % d'évènements cliniques en ITT et en PP au bénéfice systématique de la stratégie clinique par rapport à la stratégie génomique. <p>Question 3 : Utilisation de Mammaprint en désescalade si C haut risque/ G bas</p> |

⁴ Inclus à tort, non compliance au traitement alloué, erreurs d'estimation du risque génomique et/ ou clinique

⁵ Excluant toute la période d'erreurs systématiques du risque génomique du 24 mai 2009 au 30 janvier 2010 (

⁶ Stratégie guidée par la génomique = Patientes C Haut/G Bas sous HT exclusive + patientes C Bas/G Haut sous CT/HT

⁷ Stratégie guidée par la clinique = Patientes C Haut/G Bas sous CT/HT + patientes C Bas/G Haut sous HT exclusive

Analyse primaire de l'étude MINDACT :

- population « tous profils moléculaires et ganglionnaires » avec C haut/ G Bas (23 % des tests réalisées) candidates à une désescalade ;
- n= 748 patientes sous HT exclusive mais 644 patientes pour l'analyse primaire (11 % d'attrition, pas de description de la population primaire) ;
- succès si pronostic sur la SSMD à cinq ans de la cohorte C Haut /G Bas sous HT exclusive au-dessus d'un seuil théorique et arbitraire de bon pronostic (> 92 %) = objectif atteint car 94,7 % ; IC 95% [92,5 - 96,2] ;
- C Haut /G Bas dans la population d'intérêt à tumeur luminaire pN0 sous HT exclusive cf. supplément fig S5 :
 - n=350 = SSMD à 5 ans = 93,9 % IC 95% [90,6% - 96,1%] en ITT. Ce résultat franchit le seuil de bon pronostic de 92% ;
 - aucun autre résultat approprié dans la population d'intérêt pN0 (cf. les suppléments Tableau S8 et S9 qui agrègent aux patientes C haut/C Bas luminaires et pN0 sous HT exclusive respectivement des patientes C Haut/G Bas sous CT et les patientes C Bas /G bas sous HT exclusive).

Analyses secondaires de l'efficacité de la CT adjuvante et de son bénéfice chez les patientes C Haut/G Bas sur la SSM à 5 (patientes candidates à une désescalade) :

- C Haut/ G bas « tous profils moléculaires et ganglionnaires » = efficacité de la CT établie sur la SSM dans toutes les analyses (Tableau S14, Tableau 2, Tableau S5 de la publication) :
 - HR 0,71 IC [0,50 – 1,01] en ITT (p : 0,055) Δ observé⁸ = 2,8 % ;
 - HR 0,64 IC [0,43 – 0,95] en PP (p= 0,03) Δ observé = 3 % ;
 - HR 0,57 IC [0,37 – 0,87] en PPS (p=0,009) Δ observé = 4,5 %.
- C Haut /G Bas dans la population d'intérêt luminaires pN0 (cf. supplément fig S5) :
 - SSMD HR 0,80 [0,44 – 1,45] ; p = NS ; Δ observé = 1,6 % ; IC 95% non fournie mais borne supérieure estimée proche de 4 à 5% ;
 - Aucun autre résultat rapporté, notamment sur la SSM et en PP.

Valeur prédictive d'efficacité de la CT de Mammaprint : non démontrée

- Aucun test d'interaction rapporté par les auteurs mais :
- Efficacité de la CT meilleure en cas de G Bas que de G Haut dans toutes les analyses et les critères d'évaluation (effet prédictif paradoxal de Mammaprint)
- Par ex cf. Tableau 2 : HR 0,65 en SSMD – 0,64 en SSM – 0,63 en survie globale (G Bas) versus respectivement HR 0,90 – 0,74 – 0,72 (G Haut)

Données de sécurité (EIG) ou de qualité de vie non rapportées par les auteurs.

⁸ Bénéfice de la CT adjuvante à 5 ans. Borne supérieure de l'IC 95 % proche de 5 %

Annexe 7. Tableaux de données de l'étude TAILORx {Sparano, 2018 #159}

| | |
|--|--|
| Sponsors de l'étude | Étude américaine financée principalement par le NCI ⁹ mais aussi par le fabricant Genomic Health. |
| Déclarations des liens d'intérêt | Sur 30 signataires, sept co-auteurs déclarent avoir un risque majeur de conflit d'intérêt avec le fabricant Genomic Health au cours des dernières années dont le co-découvreur du test et le détenteur du brevet délivré à Genomic Health. |
| Informations générales | <p>Période de recrutement : 04/2006 – 10/2010 avec un gel de base en mars 2018 (délai minimal <u>théorique</u> de suivi de la dernière patiente incluse proche de sept ans) N=11 232 patientes éligibles et incluses pour réaliser un RS dans le cadre de TAILORx.</p> <p>Taux d'exclusion <u>pré-allocation</u> des traitements : 8,5 % (959 patientes) dont 57,5 % pour retrait de consentement.</p> <p>Caractéristiques et répartition des quatre bras de l'étude <u>post-allocation</u> des traitements (sur les 10 273 RS réalisés) :</p> <ul style="list-style-type: none"> • risque génomique faible (RS ≤10 = 15,9 %) : allocation vers HT exclusive (bras A) ; • risque génomique intermédiaire (RS 11 – 25 = 67,2%) : allocation par randomisation vers : <ul style="list-style-type: none"> ▸ HT exclusive (bras B) ; ▸ HT+CT (bras C). <p>Risque génomique élevé (RS ≥26 = 16,9 %) : allocation vers HT + CT (bras D).</p> |
| Caractéristiques de la population incluse | <ul style="list-style-type: none"> • N= 10 273 candidates à une allocation de traitement (= ITT vraie) ; • tumeurs lumineales pT1b-2 pN0 ; 18 – 75 ans ; sans antécédent d'HT ; • en situation prévisible ou susceptible de recevoir un CT adjuvante d'après le référentiel US (NCCN, 2006) : <u>situation de désescalade potentielle</u> ; <ul style="list-style-type: none"> ▸ pT1b grade 2 ou 3 ; ▸ pT1c-pT2 (sans autres précisions). <p>Population d'analyse primaire de TAILORx (= ITT modifiée).</p> <p>ITT <u>modifiée</u> car exclusion sans suivi clinique de 554 patientes <u>post-allocation</u> des traitements (5,4 %) incluses à tort ou <u>sans données cliniques</u> :</p> <ul style="list-style-type: none"> • 10 patientes du bras A → 0,6 % de l'effectif ; • 196 patientes des bras B et C → 2,8 % de l'effectif ; • 348 patientes du bras D → <u>20 %</u> de l'effectif. <p>Caractéristiques de la population d'analyse (n= 9 719) en supplément Table S1 des suppléments :</p> <ul style="list-style-type: none"> • âges ± identiques entre les 4 bras ; • tumeurs : 24% pT2 (A, B, C) vs 33 % bras D ; • grade 1 : 34% (A) ; 29 % (B-C) vs <u>7 %</u> (D) ; • grade 3 : 7% (A) ; 13-14% (B-C) vs <u>50%</u> (D) ; • RE - : 0 % (A,-B-C) vs 3% (D) ; • RP - : 2% (A) ; 8 % (B-C) vs <u>30 %</u> (D). <p>Répartition différente du risque génomique dans la population d'analyse (après 5,4% d'exclusions post-allocations).</p> <ul style="list-style-type: none"> • RS≤10 (17 %) ; • RS 11-25 (69 %) ; |

⁹ National Cancer Institut (US)

- RS \geq 26 (14 %).

Taux de patientes à bas risque clinique (selon Adjuvant ! Online v8 modifiée) = patientes sans aucune indication de CT adjuvante au vu des CCP traditionnels :

- bras A : 78 % ;
- bras B et C : 74 % et 73 % ;
- bras D : 43 %.

Les patientes à haut risque clinique (C haut) selon Adjuvant ! pouvaient avoir une indication incertaine ou clairement établie de CT adjuvante au regard des CCP traditionnels.

Description des protocoles d'HT reçus post-randomisation (bras B – C) :

- si prémenopausée :
 - Tamoxifène : \pm similaires entre B et C ;
 - suppression ovarienne : 15 % (B) vs 11 % (C) ;
 - anti-aromatase : 4% (B) vs 9% (C).
- si post ménopausée :
 - protocoles \pm similaires entre B et C ;
 - mais taux non compliance déséquilibrée : 0,5 % (B) vs 2 % (C).

Description des protocoles de CT reçus post-randomisation (Bras C) :

- Taxanes + Cyclophosphamide : 56 % ;
- Anthracycline (\pm taxanes) : 36 % ;
- CMF : 7%.

Non adhésion au traitement alloué (cross-over des traitements) :

- RS \leq 10 (bras A) : 0,5% (n=8)
- RS 11-25 (bras B) : 5,4 % (n=185)
- RS 11-25 (bras C) : 18,4 % (n= 608)
- RS \geq 26 (bras D) : 6,4 % (n=89 ; cf. Tableau S2 des suppléments)

Violations majeurs du protocole (cross-over, inclus à tort, arrêt prématuré du traitement) :

- bras A : 1,1 % ;
- bras B : 7 % ;
- bras C : 21,6 % ;
- bras D : 25,1%.

Sorties prématurées (selon la figure 1) :

- bras A : 9,7 % ;
- bras B : 11,5 % ;
- bras C : 14,3 % ;
- bras D : 22,9 %.

| | |
|--|---|
| | <p>Arrêts prématurés du traitement : non rapportés</p> |
| <p>Populations d'analyses et plan statistique</p> | <p>Essai en « surpuissance statistique » : + 57 % (6 007/4 390) du nombre des patientes randomisées (bras B-C) par rapport au nombre initial de sujets nécessaires (amendement) en raison du <u>plus faible risque de récurrence observé</u> (4 390 au départ à 6 860 patientes post-amendement pour 6 907 réellement randomisées).</p> <p>Analyses prévues pour les patientes à risque génomique intermédiaire (RS 11 – 25) = Comparaison entre le bras B et C.</p> <p>Critère primaire = survie sans maladie invasive.</p> <p>Analyse de non infériorité (NI) : seuil de non infériorité de 1,32¹⁰ mais :</p> <ul style="list-style-type: none"> • calcul du NSN pour un <u>test de supériorité</u> (risque α unilatéral de 10%, risque β : 5 % (cf. publication) • P values rapportées pour la supériorité de la CTA sur l'HTA exclusive <p>Analyse secondaire en traitement reçu (répartition déséquilibrée) :</p> <ul style="list-style-type: none"> • bras B : 2 889 (43 %) ; • bras C : 3 822 (57 %). <p>Cette analyse secondaire est plus conservatrice mais n'est plus comparable sur les critères pronostiques.</p> <p>L'analyse en per protocole recommandée en cas d'analyse de non infériorité et en particulier si nombreux écarts au protocole n'a pas été fournie par les auteurs.</p> |
| <p>Résultats</p> | <ul style="list-style-type: none"> • Médiane de suivi : SSMI¹¹ : 90 mois (7 ½ ans) ; Suivi théorique minimal > 7 ans • Selon la figure S4 des suppléments (DFS) <ul style="list-style-type: none"> ▸ patientes analysées à cinq ans \approx 80 % (SSMI) sauf <u>bras D = 35 %</u> (n= 617) de l'effectif initial de 1 737 patientes ; ▸ patientes analysées à 9 ans \approx 15 % sauf <u>bras D = 4,4 %</u> (n= 77) de l'effectif initial ; ▸ données censurées non explicitées par le flow chart ? ▸ Analyses à neuf ans et celles du bras D non recevables (faible qualité et im- précision majeure). • 836 évènements observés sur la SSMI : <ul style="list-style-type: none"> ▸ 40,4 % de récurrences tumorales ; ▸ 41 % locorégionales et 59 % métastases à distance. <p>1. Analyses primaires B vs C : p value de supériorité pour un seuil de signification modifié de 5 % bilatéral)¹² :</p> <p>Critère primaire = SSMI en ITT modifiée</p> <ul style="list-style-type: none"> • HR 1,08 ; IC 95 % [0,94 – 1,24] p = 0,26 (NS) ; p = 0,13 en unilatéral (cf. sup- plément 6^E) <ul style="list-style-type: none"> ▸ Δ 5 ans : 92,8 % (B) vs 93,1 % (C) = Δ 0,3 % sans IC 95 % ; |

¹⁰ Perte d'efficacité acceptée de 32 % sous HTA exclusive

¹¹ Survie sans maladie invasive

¹² Selon le plan statistique du calcul du NSN, le seuil de signification devrait être de 10 % en unilatéral (p value < 0,1)

- ▶ Δ 9 ans 83,3 % (B) vs 84,3 % (C) = Δ 1,0 % ; sans IC 95 %.
 - survie globale : HR 0,99 p = 0,89 (NS) :
 - ▶ délai avant une première récurrence : HR 1,11 p=0,33 (NS) ;
 - ▶ délai avant métastase à distance : HR 1,10 p=0,48 (NS).
 - 2. Analyses en traitements reçus (analyse de sensibilité plus conservatrice) : p value de supériorité à 5 % (test bilatéral) :**
 - Critère primaire = SSMI**
 - HR 1,14 ; IC 95 % [0,99 – 1,31] p=0,06 (NS) ; **p = 0,03 en unilatéral** :
 - ▶ Δ à 5 ans et IC 95 % non rapportés.
 - ▶ Δ 9 ans : 83,1 % (B) vs 84,7 % (C) = Δ 1,6 % sans IC 95%
 - Valeur pronostique du RS entre 0 – 25 chez les patientes sous HT exclusive (bras A vs B):**
 - pronostic à 5 ans (en ITT modifiée) **analyse à 9 ans non recevables car trop imprécises** :
 - ▶ SSMI: 94,0 % (A) vs 92,8 % (B) = Δ 1,2 % ;
 - ▶ Récidive tumorale : 98,8 vs 96,9 % = Δ 1,9 % ;
 - ▶ Métastase à distance : 99,3 (A) vs 98,0 (B) = Δ 1,3 % ;
 - ▶ Survie globale : 98,0 (A) vs 98,1 (B) = Δ 0,1 %.
- Aucune analyse exploratoire multivariée n'a été réalisée pour identifier la valeur ajoutée du RS au-delà des CCP traditionnels.

Annexe 8. Tableaux de données de l'étude ancillaire à PLAN B {Nitz, 2017 #22}

| | |
|---|---|
| Sponsors de l'étude | <p>Étude de financement industriel allemande (WSG). Genomic Health est l'un des sponsors de l'étude</p> |
| Déclarations des liens d'intérêt | <p>Sur les 21 auteurs, dix co-auteurs ont des liens d'intérêts financiers avec Genomic Health. Trois co-auteurs étaient salariés de Genomic Health. Assistance rédactionnelle par un consultant financé par le fabricant.</p> |
| Caractéristiques de la population incluses | <p>Population ciblée : 18 à 75 ans, cancer de stade précoce et <u>localement avancé</u> (HER2 négative) ayant une indication théorique de CTA :</p> <ul style="list-style-type: none"> • pT1-4c ; pN1-3 ou • pN0 et un facteur de mauvais pronostic (grade 2-3 ou pT2 ou < 35 ans ou RH négatif ou valeur élevée de uPA/PAI-1). <p>Recrutement : 04/2009 - 10/2011 ; Amendement en août /2009 (tumeur luminale, RS ≤ 11 pN0-1 devient <u>un critère de non inclusion dans PLAN B.</u></p> <p>Inclusion : 3 198 patientes dans PLAN B</p> <ul style="list-style-type: none"> • 85,9 % lumineales (RO+) • 14,1 % triples négatives (RO-) <p>Population d'intérêt de l'évaluation HAS est représentée par :</p> <ul style="list-style-type: none"> • 2 582 à tumeur luminale pN0-1 (164 patientes RH+ étaient pN2-N3) • dont 2 274 à tumeur luminale pN0-1 post amendement (71 % des inclusions de PLAN B chez qui le RS était décisionnel) <p>Mais :</p> <p>Taux de patientes à tumeur luminale pT1c-2 pN0-1 inconnu ?</p> <p>Randomisation : 2 449 patientes dans PLAN B (76,6 % des inclusions) :</p> <ul style="list-style-type: none"> • exclusion <u>post amendement</u> des patientes RH+ ; RS ≤ 11 et pN0-1 n= 404 patientes (12,6 % des inclusions) ; • ces patientes sous HTA exclusive sont suivies prospectivement (résultats ancillaires en dehors de PLAN B). <p>Les caractéristiques clinico-pathologiques de la population à tumeur luminale pN0-1 post amendement n'est pas décrite par les auteurs que le RS soit ≤ 11 ou ≥12. Seul détail extrait du tableau, toutes les tumeurs lumineales étaient au moins pT1c (T ≥ 1,2 cm) (cf. banque tumorale).</p> |
| Populations d'analyse et tests réalisés | <p>Dans le registre RS ≤ 11 avec désescalade effective de la CTA (n= 348 / 404) :</p> <ul style="list-style-type: none"> • un refus d'HTA exclusive a concerné 50 patientes (qui ont été randomisées pour recevoir une CTA dans PLAN B) • six patientes ont refusé tout suivi clinique (sorties d'étude). <p>Aucune description précise n'a été faite de la population luminale pN0-1 et RS ≥ 12.</p> |

| | |
|------------------|---|
| Résultats | <p>A cinq ans, la SSM sous HT exclusive était de 94,2 % [IC 95 % : 91,2 - 97,3] sans différence entre les patientes pN0 et pN1. Ce risque à cinq ans pour les patientes pN0-1 sous HT exclusive a été estimé sur environ 30 % des patientes du registre (ce qui explique l'IC 95% peu précis).</p> <p>Aucun résultat pronostique précis à cinq ans pour RS \geq 12.</p> <p>Aucun résultat ne permet d'évaluer le profil clinico-pathologique ni le risque clinique de récurrence de ces patientes.</p> <p>Aucune comparaison entre les patientes du registre RS \leq 11 d'un côté et celles ayant un RS \geq12 de l'autre à cinq ans n'était permise en raison de la non comparabilité de ces groupes sur le traitement reçu, les CCP traditionnels (non décrits) ou le statut ménopausique des patientes...</p> |
|------------------|---|

Annexe 9. Tableaux de données de l'étude OPTIMA Prelim {Bartlett, 2016 #177}

| | |
|--|--|
| Population ciblée | <p>Patientes > 40 ans à tumeur luminale ayant une indication clairement établie de CT adjuvante :</p> <ul style="list-style-type: none"> • pTx ; pN1a-2a ; • pT2 ; pN0 avec T > 3cm. <p>n= 350 patientes éligibles dont 313 randomisées.</p> |
| Schéma d'étude | <p>Étude prospective comparant la concordance décisionnelle entre trois signatures génomiques différentes : Oncotype Dx (RS ≤ 25 ou RS > 25) ; Prosigna (ROR-PT bas/intermédiaire ou haut risque) ; Mammaprint (bas ou haut risque).</p> |
| Populations d'analyse et tests réalisés | <p>N=302 tumeurs analysées OPTIMA Prelim = 96,5 % des patientes randomisées dans l'essai OPTIMA en cours de réalisation.</p> <p>Succès technique : 98,9 et 99,7 % selon le test.</p> <p>Taux de risque génomique indiquant une décision de CT adjuvante :</p> <ul style="list-style-type: none"> • 18 % des tests Oncotype Dx (RS > 25) ; • 39 % des tests Mammaprint ; • 34 % des tests Prosigna. |
| Résultats | <p>Taux de concordance à l'échelle individuelle (score kappa κ) :</p> <ul style="list-style-type: none"> • κ = 0,40 [0,30 – 0,49] entre Oncotype Dx et Mammaprint ; n=297 échantillons analysés (98,3%) ; • κ = 0,44 [0,33 – 0,54] entre Oncotype Dx et Prosigna ; n= 298 échantillons analysés (98,7%) ; • κ = 0,53 [0,43 – 0,63] entre Prosigna et Mammaprint ; n= 295 échantillons analysés (97,7%). <p>La concordance était faible [entre 0,20 – 0,40] à modéré [entre 0,40 – 0,60] entre les trois signatures.</p> <p>Modifications décisionnelles si seuil binaire (cf. Supplément Table S3) :</p> <ul style="list-style-type: none"> • Oncotype Dx vs Prosigna : 22,5% de modifications décisionnelles : <ul style="list-style-type: none"> ▸ Désescalade avec Oncotype Dx (si RS ≤ 25 : 19,8% des cas par rapport à Prosigna (n= 59) ▸ Escalade avec Oncotype Dx dans 2,7 % des cas (n=8) • Oncotype Dx vs Mammaprint : 25,6 % de modifications décisionnelles; <ul style="list-style-type: none"> ▸ Désescalade thérapeutique avec Oncotype Dx (si RS ≤ 25) : 23,6% des cas (n= 70) ▸ Escalade avec Oncotype Dx dans 2 % des cas (n= 6) • Mammaprint vs Prosigna : 21,7 % de modifications décisionnelles <ul style="list-style-type: none"> ▸ Escalade avec Mammaprint : 13,2 % des cas (n= 39) ▸ Désescalade avec Mammaprint : 8,5 % des cas (n= 25). <p>Environ une patiente sur quatre a une modification décisionnelle pour l'ensemble des SG</p> |

Annexe 10. Compte rendu de la réunion du GT du 02 mai 2018

COMPTE RENDU

Type de réunion : Première réunion du groupe de Travail

Titre : Évaluation de l'utilité clinique des signatures génomiques pour aider à la décision d'une chimiothérapie adjuvante dans le cancer du sein de stade précoce

Date : 02/05/2018

Objectifs de la réunion

Discuter, définir les éléments clés du protocole d'évaluation,

- le périmètre d'évaluation, qui intègre les questions cliniques pertinentes à traiter, la population cible, la définition du rôle et de la stratégie d'utilisation des signatures, des comparateurs et de la stratégie de prise en charge habituelle, des critères de d'évaluation cliniquement pertinents qui seront à hiérarchiser selon leur importance du point de vue des patientes ;
- les schémas des études à retenir permettant d'établir l'utilité clinique des SG et la durée de suivi minimal requis des patientes

Préambule HAS :

Une présentation générale a été faite par le chef de service du SEAP (service d'évaluation des actes professionnels) pour exposer les missions de la HAS, la méthode d'évaluation des actes professionnels, le processus de sélection des experts (analyse des liens d'intérêts en comité de validation et vérification en début de réunion), le rôle et le fonctionnement du groupe de travail (GT) dans ce cadre particulier. Une rapide présentation de l'autosaisine de la HAS a été faite aux experts reprenant les éléments qui figurent dans la feuille de route validée par le collège de la HAS et publiée sur le site internet de la HAS.

Le chef de projet a présenté le contexte médicoteknique de cette évaluation dont :

- les éléments d'appréciation de l'utilité clinique d'un nouveau marqueur guidant la sélection d'un traitement ou d'une prise en charge thérapeutique (rappel du double aspect d'impact décisionnel et clinique), que le marqueur ait une valeur de réponse pronostique, prédictive ou les deux à la fois ;
- la description des SG et l'objectif clinique à atteindre dans le cancer du sein de stade précoce ;
- le rappel du processus de sélection menant à la population source de l'évaluation (ou réservoir des patientes cibles) susceptible d'avoir un intérêt clinique et un réel impact populationnel en cas d'utilisation d'une SG ;
- la description des trois modalités d'utilisation d'une SG identifiées dans la littérature par la HAS :
 - remplacement de la stratégie décisionnelle habituelle par la décision basée sur un score génomique ;
 - intégration du score génomique dans un modèle de prédiction du risque et du bénéfice thérapeutique tenant compte des autres critères existants (critères clinico-pathologiques -CCP) ;
 - l'utilisation du score génomique en situation de décision initiale de CT adjuvante basée sur des « critères de mauvais pronostic » classiques en vue d'une désescalade thérapeutique permise par un score génomique bas.

La HAS a rappelé aux membres du GT que l'objectif de ce travail n'est pas l'élaboration d'une recommandation nationale de bonnes pratiques professionnelles. Les discussions sur le positionnement du groupe concernant les différents CCP habituellement pris en compte par les professionnels pour décider ou non d'une CT adjuvante n'ont pas vocation à entériner ou modifier les pratiques professionnelles françaises actuelles. L'objectif de cette première réunion est de définir la place et l'utilisation des SG dans une stratégie guidant la décision de CT adjuvante dans le cadre d'un **rapport d'évaluation technologique en vue d'un remboursement** en se fondant sur les dernières données acquises de la science et de la pratique française.

Préambule des experts :

Les experts veulent souligner la bonne répartition sur le territoire français des professionnels concernés (anatomopathologistes, biologistes médicaux), la qualité de leur formation initiale et continue et la grande rigueur du contrôle qualité mis en place en France. De ce fait, les professionnels s'inquiètent, à leur niveau, de l'impossibilité d'effectuer leur propre contrôle qualité concernant les SG dites centralisées¹³ (*Oncotype Dx, Mammaprint*). Il est demandé à la HAS que l'évaluation technologique prenne tout particulièrement en compte l'ensemble de ces éléments pour apprécier l'utilité clinique des SG en contexte français.

Les experts du GT soulignent que, du point de vue des professionnels français, dans le cas du cancer du sein de stade précoce, **leur besoin aujourd'hui non couvert est de permettre une optimisation de la décision de CT adjuvante en situation d'incertitude majeure par un nouveau marqueur tumoral performant**, et moins de permettre une désescalade thérapeutique chez les patientes ayant une indication clairement établie de CT adjuvante sur la base des CCP conventionnels.

1. Exactitude pronostique et prédictibilité de l'efficacité d'une CT adjuvante : intérêt de l'étude prospective et comparative « ad hoc » évaluant la pertinence thérapeutique des décisions basées sur une SG

Il apparaît que le « postulat » selon lequel l'efficacité relative de la CT adjuvante serait globalement stable¹⁴ (sur la réduction des récurrences tumorales et de la mortalité) entre les différents sous-groupes de patientes et pour les différentes molécules administrées n'est pas consensuellement admis par les membres du GT. Ce postulat s'appuie principalement sur les résultats de la méta-analyse du groupe de l'*Early Breast Cancer Trialists' Collaborative Group* (EBCTCG - Peto *et al*, Lancet 2012) repris par les experts de l'ASCO en 2016 pour leur permettre d'évaluer l'utilité clinique des marqueurs tumoraux dans le cancer précoce du sein (Harris *et al*, JCO 2016).

Néanmoins, certains experts du GT ont rappelé que l'absence de recueil du statut moléculaire HER2 et Ki67 des tumeurs incluses dans la méta-analyse de l'EBCTCG est une limite importante à ce postulat. Pour eux, il existe d'autres preuves issues des données recueillies **en contexte néoadjuvant** qui mettent indirectement en évidence l'existence de critères prédictifs d'efficacité de la CT. Néanmoins, ces conclusions se basent sur un objectif clinique différent en néoadjuvant de celui visé en adjuvant (taux de récurrence et de survie globale pour l'un ou réponse tumorale en vue de permettre un traitement conservateur pour l'autre). D'après l'expérience personnelle de certains experts et d'après les recommandations en vigueur de l'ESO-ESMO (Cardoso F *et al*, Ann Oncol, 2017), il

¹³ Notion de « boîte noire » pour certains.

¹⁴ C'est-à-dire sans critères prédictifs d'efficacité authentifiés.

existerait des critères prédictifs d'hormono-résistance (ou de chimiosensibilité¹⁵) en contexte néoadjuvant utiles pour guider le choix entre une CT et une hormonothérapie (HT).

Cette controverse qui consiste à devoir accepter (ou non) qu'il existe des critères de prédiction d'efficacité de la CT adjuvante met en exergue pour une majorité des membres du GT, l'intérêt de ne pas réduire l'évaluation de l'utilité clinique d'une SG à ses performances pronostiques (exactitude pronostique). Selon une majorité des experts, l'utilité clinique d'une décision basée sur une SG doit faire la preuve directe d'un **impact clinique favorable** sur l'état de santé des patientes (balance bénéfique/sécurité positive)

La validation d'une SG permettant d'optimiser la décision de CT adjuvante (ayant fait par ailleurs la preuve de son efficacité globale) passe par une **étude prospective et comparative évaluant l'impact décisionnel et clinique des décisions prises par la génomique tumorale sur des critères cliniques pertinents**, notamment si :

- la balance bénéfice sécurité est comparée entre la stratégie décisionnelle basée sur la génomique et la stratégie habituelle (essai prospectif de stratégie) ;
- la balance bénéfice sécurité est comparée entre chaque catégorie de risque génomique, la SG étant réalisée chez toutes les patientes à l'inclusion (essai prospectif stratifié sur le score génomique).

2. Prérequis et définition des populations éligibles à une prescription de SG – établissement des critères de mauvais pronostic et de susceptibilité à recevoir une CT adjuvante

Conditions de prescription d'une SG

Un rappel a été fait par le GT concernant la pertinence de ces prescriptions. Afin d'évaluer la catégorie de risque de récurrence et, par voie de conséquence, de décider de prescrire une SG, il est nécessaire pour les cliniciens de disposer en réunion de concertation pluridisciplinaire (RCP) de l'ensemble des CCP habituels. Par ailleurs, dans le cadre d'une **décision médicale partagée**, informer correctement la patiente de son risque *a priori* de récurrence et recueillir ses préférences concernant une administration de CT adjuvante sont des éléments indispensables qui peuvent clairement contre indiquer la prescription de SG (non adhésion anticipée de la patiente à toute décision extérieure à son choix). Le recueil de l'**accord d'une patiente bien informée** serait également indispensable. En l'absence d'un référentiel commun au niveau national, les membres du GT ont rappelé **la grande hétérogénéité en France des décisions de CT adjuvante** prises en RCP selon les centres. Certains ont émis l'hypothèse de recourir pour la prescription d'une SG à des « RCP régionales » du réseau de soins en cancérologie pour favoriser cette harmonisation.

Trois situations cliniques

Après discussion, le GT s'est positionné sur les trois situations cliniques distinctes d'utilisation de SG proposées par la HAS.

Les trois situations sont une des utilisations suivantes :

- ▶ l'ensemble de la population source à risque intermédiaire, c'est-à-dire pT1c-2¹⁶, pN0-1 de nature moléculaire luminale (RH+, HER2+), très proche des recommandations françaises de Saint Paul de Vence 2015 ;

¹⁵ Grade histo pronostique, RE et RP, HER2, Ki67, statut ménopausique...

¹⁶ Les tumeurs multifocales sont acceptées et les tumeurs canalaire sont plus fréquentes et à privilégier.

- ▶ le sous-groupe au sein de cette population source des cas équivoques en incertitude décisionnelle majeure ;
- ▶ le sous-groupe des patientes au sein de cette population source susceptible (voire en situation fortement prévisible) de recevoir une CT adjuvante dans la pratique courante.

Cas des tumeurs pT1b luminales de grade 3 :

Pour certains experts, l'inclusion de **certaines tumeurs infracentimétriques pT1b luminales** ayant des facteurs mauvais pronostiques (en particulier un grade 3) serait également une indication envisageable de SG (discordance clinico-pathologique). Certains experts ont souligné en effet la réelle incertitude du clinicien pour ces petites tumeurs rarement incluses dans les essais thérapeutiques et dont la prise en charge est peu standardisée, au cas par cas (EURISTIC - Spielmann et al, 2016). Néanmoins, Il est rappelé que ces cas sont plutôt rares. En effet, les résultats de l'étude prospective française ODISEE (Dalenc et al, 2017) ont montré que pour les 90% de l'effectif qui étaient des tumeurs pT1ab pN0 de nature luminaire, **moins de 5 % étaient de grade 3** et environ 3 % possédaient des embolies vasculaires. Chez ces patientes, **la CT adjuvante était prescrite dans seulement 2 % des cas**. De même, les tumeurs luminales pT1b pN0 présentaient globalement un **excellent pronostic** (0 métastase à distance, 2 récurrences locorégionales sur un suivi clinique de 5 ans d'environ 500 patientes).

Compte tenu du pronostic très favorable de ces tumeurs pT1b luminales représentant presque 15 à 20 % des cancers du sein (EURISTIC- Spielmann et al, 2016), l'intérêt d'inclure les tumeurs pT1b luminales dans la population source n'est pas paru opportun à la HAS du fait de la faible proportion de patientes présentant des critères de mauvais pronostic et de l'incertitude autour de leur devenir clinique. La HAS considère qu'il faudrait préalablement considérer le pronostic de ces petites tumeurs luminales de grade 3 par une étude pronostique bien conduite.

Critères de mauvais pronostic rendant la situation susceptible ou prévisible d'une décision de CT adjuvante en RCP

Il a été admis par le GT que dans la population source proposée (40 % environ des cancers non métastatiques selon les estimations du GT), il existe des **critères binaires** de mauvais pronostic¹⁷ clairement soutenus par la littérature :

- ▶ Patiente jeune < 35/40 ans ;
- ▶ Taille tumorale entre 2 et 5 cm (pT2) ;
- ▶ Grade 3 histo-pronostique ;
- ▶ Envahissement ganglionnaire (pN1), avec au moins une macrométastase¹⁸ ;
- ▶ Présence d'embolies lymphovasculaires.

Pour le GT, **un seul de ces critères** chez une patiente de la population source rendrait la patiente susceptible de recevoir une CT adjuvante dans certains centres (notamment si grade 3 ou si pN1 avec macrométastases), mais pas dans d'autres. L'existence **d'au moins deux de ces critères** chez une patiente de la population source rend la situation très prévisible d'une décision de CT adjuvante partout en France.

¹⁷ D'autres critères quantitatifs sans seuil précis (plus subjectifs) sont l'expression des récepteurs à la progestérone et aux œstrogènes en cas de RH positifs

¹⁸ D'après les experts, la découverte d'une ou plusieurs micrométastases ganglionnaires axillaires a actuellement peu de conséquences pratiques sur la prise en charge des patientes (pas de reprise chirurgicale, pas de curage axillaire, pas d'extension du champ d'irradiation autour du sein, pas d'indication à une CT adjuvante). Ces métastases de petite taille étaient classées pN0 par le passé.

Pour la HAS, les deux situations cliniques ci-dessus (présence d'un ou d'au moins deux critères de mauvais pronostic) pourront être appréhendées, soit conjointement soit séparément, avec un objectif de désescalade thérapeutique basée sur le score génomique à la différence des situations dites en incertitude décisionnelle majeure dont l'objectif clinique sera l'optimisation de la décision en fonction du score génomique.

Cas particulier du marqueur Ki67

L'utilité clinique du marqueur tumoral Ki67 n'a pas été admise pour tous les membres du GT. Bien que son financement soit actuellement inclus dans un acte CCAM non spécifique, son dosage n'est à ce jour ni recommandé par la HAS (guide ALD, 2010), ni évalué par elle. Sa valeur pronostique semble admise mais **l'absence de seuil décisionnel valide associé à une catégorie intermédiaire particulièrement large (score situé entre 10 et 30 % selon le GT)** et son manque de reproductibilité entre les équipes ne permet pas de l'inclure avec confiance dans la liste des « critères binaires de mauvais pronostic » pour cette évaluation. Le score quantitatif du Ki67 (valeur continue) est néanmoins intégré dans les principaux modèles de prédiction du risque de récurrence et de bénéfice thérapeutique pouvant être utilisés comme un comparateur des SG pour cette évaluation. Ce marqueur de prolifération est à ce jour un **comparateur direct des SG** (au sein des dosages IHC 4) pour la décision de CT adjuvante et reste en cours d'évaluation par la communauté scientifique (Recommandations de l'ASCO sur les biomarqueurs en 2016 ; Penault-Llorca et al, 2017).

3. Discussion autour des modalités d'utilisation d'une SG dans la population source (Q1) : la substitution systématique de la stratégie habituelle par une décision génomique est-elle envisageable ?

Cette question d'évaluation cherche à savoir si une SG peut remplacer les CCP classiques pour améliorer la décision de CT adjuvante. Cette question ne peut être tranchée que si le devenir des patientes dont la décision est basée sur les CCP classiques est comparé à celles dont la décision est basée sur le score génomique (analyse des cas discordants). Pour les experts, il ne devrait pas être question de légitimer le remplacement complet des CCP classiques au profit exclusif des SG. En effet, les CCP permettent dans un premier temps de caractériser la population source éligible à une SG.

Dans le cadre de cette évaluation, la question sera de savoir si dans cette population à risque intermédiaire telle que définie par les CCP, la décision de CT adjuvante peut être exclusivement basée sur le score génomique.

4. Discussion autour des critères d'évaluation pertinents et de leur hiérarchisation

Survie globale

Le GT s'est positionné sur le fait que l'amélioration de la **survie globale** est l'objectif « théorique » à atteindre en oncologie mais qu'il est particulièrement difficile à mettre en évidence en contexte adjuvant. Chez ces patientes à risque intermédiaire de récurrence, elle nécessiterait en pratique un nombre très important de sujets et un suivi clinique particulièrement long.

Survie sans métastase à distance

De ce fait, la discussion s'est focalisée autour de la pertinence de « la survie sans métastase à distance » par rapport à la « survie sans maladie invasive/récurrence » pour évaluer l'utilité clinique des SG. D'après certains membres du GT, la CT adjuvante n'aurait pas d'effet sur les récurrences locales et son efficacité serait controversée sur les récurrences régionales (pour rappel, environ 30 % des premières récurrences sont locorégionales dans le cancer du sein). Ce postulat de non efficacité de la CT adjuvante sur les récurrences locorégionales n'est pourtant pas soutenu par les résultats de l'étude NSAPB 20 - comparant HT adjuvante versus HT+CT adjuvante chez des patientes RH+,

pN0 et ayant reçu en cas de traitement conservateur de la radiothérapie adjuvante (Fisher et al, JNCI 1997).

D'un autre côté, la survie sans métastase à distance (ou le délai avant une première métastase à distance) a été particulièrement utilisée durant la phase de développement et de validation des SG dans le cancer du sein. Selon les experts du GT, l'apparition d'une métastase à distance est l'évènement pronostique et clinique le plus grave dans l'histoire de la maladie cancéreuse. Il a été souligné cependant l'importance de bien connaître les modalités d'estimation de la survie sans métastase à distance dans les études pour apprécier sa pertinence clinique au cas par cas. Par exemple, ce critère composite peut conduire à censurer de l'analyse ou à ignorer les patientes ayant eu une première récurrence locorégionale ou un second cancer invasif (récurrence locorégionale pourtant pertinente du point de vue d'une patiente et corrélé fortement à la survenue d'une métastase à distance ultérieurement). Ces censures dites « informatives » présentent un risque de biaiser l'estimation rapportée. De même, le modèle d'estimation utilisé pour la survie sans métastase à distance doit être adapté (l'analyse de Kaplan-Meier n'est plus légitime) pour prendre en compte les évènements compétitifs de la métastase à distance. Pour toutes ces raisons, **la survie sans métastase à distance** reste un **critère secondaire** des études thérapeutiques en oncologie qui ne sont jamais construites autour de celui-ci (besoin de beaucoup plus de suivi et de patientes), que ce soit pour des raisons cliniques, méthodologiques et statistiques.

Survie sans maladie invasive (ou sans rechute)

Il est rapporté par les membres du GT que la « **survie sans maladie invasive** »¹⁹ ou « **sans rechute** » est le **critère de jugement principal des essais thérapeutiques** qui ont permis d'évaluer la balance bénéfice/sécurité de la CT adjuvante pour obtenir une nouvelle indication ou pour autoriser un nouveau protocole de CT adjuvante ou de thérapie ciblée (HT ou antiHER2) dans le cancer du sein de stade précoce. Ce critère composite prend en compte l'ensemble des évènements importants pour établir une balance bénéfice sécurité, à l'exception des EIG de grade 3 et 4. À ce titre, il permet de comptabiliser plus d'évènements cliniques pertinents (du point de vue de la patiente) que la survie sans métastase à distance (qui conduit en parallèle à une puissance statistique également plus élevée). En effet, du point de vue de la patiente qui espère une guérison de sa maladie cancéreuse, toute première récurrence (ou tout second cancer invasif) impacte négativement, **le service rendu de la CT adjuvante reçu ; ce protocole reçu se révélant de son point de vue inutile**. L'ensemble des groupes d'experts du domaine²⁰ (DATECAN , STEEP , ICHOM) et les guides destinés aux industriels élaborés par l'agence européenne et américaine du médicament (EMA -FDA) comme ceux d'autres agences de remboursement que la HAS, appellent tous à utiliser la **survie sans maladie (ou sans rechute)** pour juger de l'impact thérapeutique d'une CT adjuvante ou de la surveillance de la maladie cancéreuse. Pour les membres du GT, **la survie globale et la survie sans maladie invasive (ou la survie sans rechute) sont des critères susceptibles d'être des critères primaires d'évaluation en contexte adjuvant**.

Qualité de vie

Pour eux, la modification de la qualité de vie n'est pas suffisamment adaptée à cette évaluation pour en faire un critère primaire (facteurs de confusion, moment d'analyse non défini, manque de sensibilité des questionnaires).

Toxicité

L'intégration des **évènements graves liés à la toxicité de la chimiothérapie est recommandée par le GT en critère primaire et tout particulièrement les complications différées ou à long terme** (cancers et leucémies chimio induits, insuffisance cardiaque de grade 3 et 4). Il semble que les complications aiguës soient moins fréquentes que par le passé et mieux accompagnées par les

¹⁹ Incluant les seconds cancers qui peuvent être difficilement distinguables d'une métastase à distance.

²⁰ qui a pu inclure des patientes pour l'ICHOM

oncologiques (prévention de la neutropénie fébrile notamment). Leur impact clinique est variable car quelque fois sans gravité réelle pour les patientes.

Suivi clinique minimal

Le GT se positionne en faveur d'une analyse des événements cliniques comme les récurrences tumorales dans le cadre **d'un suivi clinique minimal de cinq ans des patientes**. Du point de vue des experts, le bénéfice clinique de la CT adjuvante s'appuie sur le caractère prolifératif des cellules tumorales et de ce fait est maximal dans les cinq premières années du suivi (avec un pic de récurrences autour de 2,5 ans de suivi régulièrement constaté dans les études).

5. Discussion autour de la catégorie de risque génomique « intermédiaire » pour certaines SG (Oncotype Dx, Prosigna)

Selon le GT, l'existence d'une catégorie de « risque génomique intermédiaire » diminue la confiance des cliniciens. Notamment en situation d'incertitude décisionnelle (zone grise), cette catégorie « intermédiaire » diminue l'impact décisionnel de ces SG à hauteur de la proportion de patientes concernées par cette catégorie, et par voie de conséquence, leur impact thérapeutique et clinique par rapport à la stratégie habituelle. Une SG classant la patiente selon deux catégories de risque génomique répond à une réelle exigence pratique du point de vue des professionnels (traiter ou ne pas traiter selon un seuil décisionnel établi par le fabricant).

6. Discussion autour de l'utilisation en France des modèles en ligne de prédiction du risque et du bénéfice thérapeutique attendu de la CT adjuvante

Les experts du GT soulignent que l'information génomique contenue dans une signature devrait être dans l'idéal combinée aux informations apportées respectivement par les CCP classiques. En revanche, en ce qui concerne les outils multicritères d'aide à la décision de CT adjuvante, l'INCa ne les recommandait pas en 2012, bien qu'ils puissent être en théorie les seuls outils, une fois validés, capables d'intégrer l'ensemble des informations contenues dans les CCP recueillis en routine. D'après les experts interrogés, ces outils de prédiction du risque et du bénéfice attendu des traitements systémiques seraient de fait très peu utilisés en France. L'outil Adjuvant ! Online est d'ailleurs devenu inaccessible en pratique. L'outil PREDICT V2.0 est peu connu des professionnels français.

7. Notion de seuil de risque décisionnel et de bénéfice absolu attendu

Dans le cadre d'un modèle de prédiction (cf. thème n° 6 ci-dessus), un seuil décisionnel de mortalité globale à dix ans d'au moins 10 % a été proposé pour pouvoir prescrire une CT adjuvante. En accord avec la littérature disponible, les experts du GT se positionnent sur l'absence d'intérêt clinique de la CT adjuvante en cas de bénéfice absolu prédit de moins de 3 % sur la survie sans récurrence à cause de sa toxicité importante. Entre 3 et 5 % de bénéfice théorique prédit, le GT recommande le recours à une **décision médicale partagée avec la patiente** pour décider d'une CT adjuvante ou non. Au-delà de 5 % de bénéfice absolu prédit pour les patientes, la CT adjuvante serait clairement indiquée.

Annexe 11. Compte rendu de la réunion du GT du 02 juillet 2018

COMPTE RENDU

Type de réunion : Seconde réunion du groupe de travail (GT)

Titre : Évaluation de l'utilité clinique des signatures génomiques (SG) pour aider à la décision d'une chimiothérapie adjuvante dans le cancer du sein de stade précoce

Date : 02/07/2018

Objectifs de la réunion

Discuter de manière collégiale et argumentée des résultats des études retenues par la HAS pour répondre aux questions d'évaluation relatives à l'utilité clinique des SG.

Proposer des éléments de réflexion autour des perspectives envisageables d'utilisation des SG en contexte français.

Préambule HAS/Experts :

En raison des nombreuses études portant sur la validation clinique des SG et de l'intégration de déterminants propres aux critères clinico-pathologiques (CCP) traditionnels dans leur algorithme de calcul, les experts leur reconnaissent une **valeur pronostique intrinsèque**.

Néanmoins comme rappelé dans le rapport de l'INCa, la question principale est plutôt d'apprécier leur utilité clinique en démontrant une **amélioration de l'état de santé** chez les patientes chez qui les SG ont été utilisées. Les SG devraient ainsi faire la preuve de leur **valeur ajoutée** par rapport aux CCP traditionnels pronostiques et/ou prédictifs d'efficacité.

Concernant le modèle de prédiction du risque **Adjuvant ! Version 8 (avec HER2)** utilisé pour classer sur la base des CCP les patientes en bas ou en haut risque de récurrence dans l'étude MINDACT et TAILORx²¹, les experts du GT signalent que cet algorithme décisionnel n'est plus, à leur connaissance, utilisé actuellement dans les réunions de concertation pluridisciplinaires (RCP) françaises. Ils estiment par exemple qu'il peut surestimer les indications prévisibles de CT adjuvante pour les grosses tumeurs (pT2) de plus de 3 cm sans critère d'agressivité biologique (grade 1).

Le principal reproche qui est fait à ce modèle est l'absence de reconnaissance d'une « zone d'incertitude décisionnelle majeure » qui représente en fait une partie importante de la population d'évaluation « source » de la HAS (cf. question d'évaluation n° 2). Aujourd'hui ce sont plutôt les référentiels de bonne pratique qui prévalent en France.

Pour estimer l'utilité clinique d'une SG dans le cancer du sein luminal de stade précoce, il a été rappelé aux experts les dénominations suivantes :

- **le risque de base** (pronostic de la patiente) est le risque de faire une récurrence sous hormonothérapie (HT) exclusive à 5 ou 10 ans (probabilité de survie ou de récurrence en %) ;
- **l'efficacité relative** d'une CT adjuvante est la réduction relative du risque de faire une récurrence sous CT + HT par rapport à l'HT exclusive (exprimée en *hazard ratio* ou HR²²) ;

²¹ Et pour établir les critères d'éligibilité de l'étude randomisée d'utilité clinique OPTIMA en cours de réalisation avec le test Prosigna.

²² Le HR est proche du risque relatif de l'effet d'un nouveau traitement au cours du temps par rapport à un traitement de référence en présence de données censurées (= analyse de survie). Par exemple, un HR de 1,32 en thérapeutique signifie qu'il existe une augmentation de 32 % de la proportion de récurrences observées par rapport au traitement de référence (effet délétère du nouveau traitement). À l'inverse, un HR observé de 0,80 signifie une diminution de la propor-

- la **réduction du risque absolu de récurrence** (exprimée en %) **qui est le bénéfice clinique du traitement** (est égale à la perte de chances à accepter ou non en cas de décision de désescalade thérapeutique), mesure utile pour déterminer l'intérêt thérapeutique réelle d'une nouvelle stratégie.

Ainsi l'intérêt d'une SG **en situation adjuvante** dans le cancer localisé du sein se résume à l'équation clinique suivante :

Risque de récurrence sous HT (en %) x Quantité d'effet relatif de la CT adjuvante sur les récurrences (HR, sans aucune unité) = **bénéfice absolu de la CT adjuvante** (en Δ %) dans une population donnée²³ = **utilité clinique**

1- Analyse et discussion des résultats des études

Une valeur pronostique mais pas de démonstration de l'utilité clinique ni de valeur prédictive d'efficacité de la CT adjuvante au sein de la population source²⁴ de l'évaluation de la HAS

Concernant les questions d'évaluation n° 1, n° 3 et n°4 du rapport provisoire, les résultats des études répondant à ces questions ont été discutés en séance. Certains experts (dont des méthodologistes spécialisés en oncologie) ont souligné l'étonnante complexité des publications retenues dans le domaine des SG et l'inhabituelle difficulté à lire, à comprendre et à interpréter rapidement leurs résultats. Ces constatations ont été renforcées par certains choix discutables de notification plutôt que d'autres concernant les résultats présentés ou d'autres éléments nécessaires à une appropriation rapide de ces nouvelles connaissances.

Pour les deux grands essais partiellement randomisés²⁵ MINDACT (Mammaprint) et TAILORx (Oncotype Dx), **les résultats ne permettent pas de démontrer l'utilité clinique de Mammaprint et d'Oncotype Dx dans la population de l'évaluation HAS** et de modifier avec une confiance suffisante les pratiques actuelles basées sur les CCP traditionnels dans ce cas (*cf.* plus haut).

Les raisons avancées par les experts et qui corroborent par ailleurs l'analyse de la littérature réalisée par la HAS sont les suivantes :

- Les **populations incluses**²⁶ dans les deux études ne sont pas celles attendues par les membres du GT pour l'utilisation d'une SG en 2018 (c'est-à-dire les patientes à risque intermédiaire en situation d'incertitude décisionnelle majeure) ;
- L'existence durant la conduite des essais d'un taux important d'**écarts majeurs au protocole** (20 à 35 % selon les groupes de cas discordants randomisés dans MINDACT ; 22 % dans le bras C sous CT (score 11 à 25) et 25% dans le bras D sous CT (score \geq 26) dans TAILORx) ;
- Les écarts fréquents au protocole dans le cadre d'un raisonnement privilégié **de non infériorité** en contexte de désescalade thérapeutique justifient de retenir prioritairement les analyses faites chez les patientes ayant **réellement reçues le traitement alloué par le protocole** (analyse en « per protocole » ou en traitement reçu « *as treated* ») au détriment de celles réalisées en intention de traiter (ITT). Cet aspect est confirmé par les exigences internationales de rédaction de la

tion de récurrences de 20 % avec le nouveau traitement par rapport au traitement de référence (effet bénéfique du nouveau traitement).

²³ Bénéfice à apprécier en tenant compte des événements indésirables graves liés à la CT adjuvante.

²⁴ Tumeurs luminales pT1c-pT2 ; pN0-N1mic-N1a.

²⁵ Randomisation des traitements limitée aux patientes à risque clinico-génomique discordant (MINDACT) et aux patientes ayant un score génomique intermédiaire compris entre 11 et 25 (TAILORx).

²⁶ TAILORx \approx 75 % patientes ayant un score entre 0 et 25 n'avaient aucune indication à une CT adjuvante selon Adjuvant ! Version 8 avec HER2, élément confirmé par les experts du groupe (= non indication dans ce cas à une désescalade donc à une SG pour désescalade) ; MINDACT : 20 % de tumeurs non luminales à haut risque de récurrence (= non indication à une désescalade ; forte corrélation avec le score génomique).

grille CONSORT pour les essais de non infériorité²⁷. Cependant, les conclusions des auteurs des études ont principalement été établies sur les résultats en ITT.

- Le taux important²⁸ (ou non rapporté²⁹) de **suivis cliniques incomplets** au cours de ces études (exclusion ou sortie précoce). Dans TAILORx, le nombre de sorties d'étude pourrait être encore sous-estimé par les auteurs au vu de la faible proportion de patientes réellement analysées à cinq ans (81 % des patientes randomisées) et surtout à neuf ans (15 % des patientes randomisées) alors que le **délai minimal de suivi** de la dernière patiente incluse devrait être normalement proche de sept ans³⁰
- Les essais ont été construits et conduits **en ouvert** avec des traitements adjuvants systémiques dont la nature étaient laissée **à l'appréciation de l'investigateur**³¹ sans détail des **arrêts prématurés de traitements**. Ces éléments exposent à des **risques de biais** de sélection dû aux croisements des traitements reçus (cross-over), de suivi, d'attrition et d'évaluation car les investigateurs, comme les patientes elles-mêmes, connaissaient tout au long du suivi leurs scores génomiques ;
- Il n'a pas été précisé si le fabricant était **en insu** des CCP traditionnels au moment de rendre chaque score génomique ;
- Les études ne rapportent ni la **qualité de vie** ni les **événements indésirables graves** liés à la CT adjuvante pour quantifier concrètement les avantages à court terme d'une désescalade.

Autres limites spécifiques à TAILORx (Oncotype Dx) relevées et discutées par le GT :

- Plusieurs experts ont souligné d'une part le caractère inhabituel du plan d'analyse statistique proposé dans le protocole pour un essai de non infériorité (risque α de 10% pour un test unilatéral) et d'autre part l'inadéquation entre ces analyses prévues et les résultats finalement présentés dans l'article original (intervalle de confiance bilatéral à 95%, les « *p values* » rapportées sont celles d'une différence entre la CT et l'absence de CT dans le cadre de tests de supériorité plutôt que les « *p values* » de non infériorité vis-à-vis du seuil clinique prédéfini de 1,32). D'autres résultats et informations importants se trouvaient malheureusement dans les suppléments mis en annexe de la publication originale.
- **Les patientes incluses étaient très majoritairement à faible risque de récurrence (ou à risque intermédiaire) sans critères suffisants justifiant une CT adjuvante en contexte français**³² ; élément appuyé par l'algorithme décisionnel ADJUVANT !³³. Dans cette population à très bon pronostic, le gain absolu de la CT adjuvante ne pouvait être cliniquement pertinent chez ces patientes (notion de futilité thérapeutique) même en situation d'efficacité thérapeutique conservée de la CT (i.e. si l'intervalle de confiance du HR observé dans l'étude franchissait le seuil de non infériorité de 1,32) ;
- Les experts ont constaté **l'absence d'analyse du sous-groupe** à risque intermédiaire avec critères de mauvais pronostic, plus proche de la population d'évaluation de la HAS (\approx 25% de l'effectif) qui aurait pu être séparée des patientes à bas risque clinique sans indication évidente de CT adjuvante (\approx 75% de l'effectif). Les résultats de TAILORx fusionnent de manière non informative ces deux populations pourtant différentes pour les professionnels.

²⁷ Piaggio G et al. Reporting of noninferiority and equivalence randomized trials JAMA : 308, 2012.

²⁸ Attrition d'environ 14,3 % des patientes randomisées dans le bras C sous CT de TAILORx.

²⁹ Cf. MINDACT.

³⁰ Délai approximatif entre la dernière inclusion en octobre 2010 et la date de gel de la base de données en mars 2018 (89 mois).

³¹ Différents protocoles de CT et d'HT administrés, quelques fois peu adaptés aux spécificités des patientes (par ex : anti-aromatase en préménopause) ; durées trop longues de l'HT...

³² Une augmentation du nombre de sujets nécessaires de 57 % a été réalisée notamment pour cette raison. La **survie sans maladie (SSM)** sous HT avait été initialement sous-estimée à 87 % à cinq ans au début du recrutement. Les SSM réellement observées à cinq ans chez les patientes incluses ont été de 93 / 94 % (+ 6 à 7 points d'amélioration du pronostic initial). Ce risque de base sous HT a été approximativement identique pour des scores compris entre 0 et 10 et entre 11 et 25.

³³ En situation de désescalade thérapeutique (score génomique compris entre 0 et 25), seulement 22 à 27 % des patientes avaient une indication potentielle de CT adjuvante selon l'algorithme Adjuvant !

- Le **risque de confusion** autour des **changements répétés de seuil d'intervention thérapeutique** au cours du développement d'Oncotype Dx³⁴ et des **deux nouveaux seuils d'intervention** proposés par les auteurs pour la prise de décision (score ≥ 26 pour les patientes de plus de 50 ans ; score ≥ 16 pour les patientes de moins de 50 ans)
- Ces nouveaux seuils ont été établis en raison du pronostic globalement péjoratif des patientes au score ≥ 26 suivies dans un registre déclaratif (bras D) et d'analyses de sous-groupes basées sur l'âge.
- Le mauvais pronostic des patientes du bras D ne peut être généralisable avec confiance à toutes les patientes de l'essai possédant un score ≥ 26 car ce registre (bras D) n'était pas représentatif de la population de TAILORx (registre en ouvert à déclaration volontaire non financé par le promoteur avec sélection et exclusion initiale de 20% des patientes sans description de leurs profils clinico-pathologiques, taux d'attrition totale de 25 % selon le diagramme de flux (*flow chart*) avec une non-compliance à la CT dans 6,4 % des cas des patientes restantes).
- Dans le bras D (score ≥ 26), il existait un risque important d'avoir sous-estimé le nombre de patientes sorties de l'étude au vu du faible nombre de patientes réellement analysées à 5 ans (**36% des patientes assignés**) et à 9 ans (**4 % des patientes assignées**) alors que seulement 189 événements avaient été observés avec un **délai minimal de suivi** des patientes normalement pourtant proche de 7 ans (cf. date de fin des inclusions).
- Les experts ont constaté qu'il y a eu probablement **une meilleure observance des traitements en cas de score génomique élevé** (en particulier compris entre 21 et 25) avec moins de sorties d'étude annoncées par les auteurs dans les suppléments de la publication. Cet aspect pourrait expliquer la conservation de l'effet thérapeutique de la CT exclusivement dans ce sous-groupe (HR = 1,32) par rapport aux deux autres sous-groupes aux scores compris entre 11 et 20 (HR $\approx 1,0$).
- La **démonstration de la valeur prédictive d'Oncotype Dx n'a pas été envisagée à la conception du design de TAILORx** car aucune randomisation concernant l'allocation des traitements n'a été prévue dans les groupes de scores compris entre 0-10 et entre 26-100. Aucun test statistique d'interaction et aucune prédictivité n'ont été rapportés et revendiqués par les auteurs entre les trois sous-groupes de scores intermédiaires compris entre 11 et 25 (*d'ailleurs la stratification de ces sous-groupes³⁵ n'a été réalisée qu'en cours de recrutement*).

Autres limites spécifiques à MINDACT (Mammaprint) relevées et discutées par le GT:

- Le profil des patientes de MINDACT candidates à une désescalade (CCP haut risque/SG bas risque) était de moins bon pronostic que dans TAILORx³⁶. Cette constatation a été confirmée au regard des SSM à cinq ans entre les deux études : MINDACT = 90 % (qui concernait des patientes pN0 et pN1 à tumeur luminale et non luminale) contre TAILORx = 93 à 94 % (qui concernait exclusivement des patientes pN0 à tumeur luminale) ;
- Les données de comparaison entre les deux stratégies (clinique ou génomique) ne permettent pas de valider Mammaprint en utilisation exclusive pour la prise de décision (Cf. Q 1) ;
- La nature des traitements systémiques reçus par les patientes (HT \pm CT) n'a pas été rapportée par les auteurs ;
- L'essai MINDACT et son objectif primaire n'ont pas été établis pour répondre directement à la question de l'utilité clinique de Mammaprint en vue d'une prise de décision thérapeutique (escalade ou désescalade) dans la population de l'évaluation de la HAS (tumeurs lumineuses) et sur la base d'analyses comparatives de non infériorité entre les deux stratégies ;
- L'objectif primaire de MINDACT reposait sur la constatation chez les patientes « tous profils moléculaires confondus » du maintien d'un bon pronostic en cas de CCP à haut risque mais en désescalade de CT sur la base d'une SG à bas risque. Le bon pronostic global de ces patientes était décrété sur la base d'une comparaison indirecte à **simple valeur théorique de 92 % de la**

³⁴ Les seuils validés et proposés actuellement par le fabricant aux cliniciens sont fixés à 18 et 31, or les seuils de TAILORx sont fixés à 10 et 26 et pour PLAN B les seuils sont encore différents...

³⁵ Scores intermédiaires compris entre 11 à 15, entre 16 à 20 et entre 21 à 25.

³⁶ En effet, toutes les patientes en situation de désescalade dans MINDACT étaient à haut risque clinique selon le modèle Adjuvant ! Version 8 avec HER2 contre seulement 25 % dans TAILORx.

survie sans métastase à distance (SSMD) à cinq ans³⁷ fixée par les auteurs jugée non acceptable selon les experts ;

- La perte de chances enregistrée chez les patientes à haut risque clinique en désescalade pouvait être très importante sur la survie sans maladie (SSM) ou la SSMD à cinq ans. Chez les patientes à haut risque clinique, un risque d'augmentation de deux à deux fois et demi le risque de récurrence tumorale en situation de désescalade a été largement appuyé par certains experts³⁸ ;
- Dans la population d'évaluation de la HAS (de moins bon pronostic également que dans TAILORx³⁹), une seule comparaison a été rapportée par les auteurs chez les patientes CCP haut risque/SG bas risque en désescalade. L'analyse de la SSMD à cinq ans en ITT est très peu informative selon les experts. Aucune conclusion ne peut être faite sur la base de ce résultat imprécis (peu d'événements observés) et incomplet (pas d'analyse complémentaire en per protocole) ;
- Aucune conclusion ne peut être faite concernant le groupe en escalade thérapeutique (CCP bas risque / SG haut risque) en raison du très faible nombre de patientes incluses dans ce sous-groupe ;
- Aucune valeur prédictive d'efficacité à la CT n'a pu être constatée avec Mammaprint malgré un design d'étude adapté (double randomisation permettant un test statistique d'interaction). Cette valeur prédictive de Mammaprint apparaît improbable au groupe d'experts au vu des résultats rapportés.

L'étude observationnelle ancillaire de PLAN B avait également pour objectif d'évaluer l'utilité clinique d'Oncotype Dx. Cependant le schéma de l'étude (non randomisée sans ajustements statistiques prévus par les auteurs) ne permet pas de déterminer la valeur ajoutée et l'utilité clinique d'Oncotype Dx. Les caractéristiques clinico-pathologiques des patientes incluses dans l'étude ancillaire et leurs données de suivi clinique sont très mal rapportées par les auteurs ce qui empêche toute transposabilité clinique vers une population bien définie pour la pratique.

Concernant la question 4 relative à l'analyse de concordance entre les différentes SG, **l'étude préliminaire de concordance OPTIMA Prelim** jugée de qualité par les experts a inclus des patientes à risque élevé et intermédiaire de récurrence. Le taux de discordance entre les trois SG analysées est de l'ordre de 25 % (concordance faible à modérée). Le groupe d'experts convient que ce taux de discordance décisionnelle important pourrait se voir augmenter dans la population d'évaluation ciblée par la HAS en situation d'incertitude décisionnelle majeure (zone grise).

Pour information, **l'étude récente de Sestak et al publiée dans JAMA Oncology en 2018**⁴⁰ a été évoquée par certains membres du GT. De l'avis général, cette étude prospective-rétrospective comparant les valeurs pronostiques de plusieurs SG ne permet pas de répondre à la question que s'est fixée la HAS en termes d'utilité clinique et en vue d'un remboursement. La concordance décisionnelle entre les différentes SG étudiées n'est pas rapportée par les auteurs. Pour ces raisons et d'autres encore⁴¹, cette étude ne peut être retenue dans le rapport d'évaluation de la HAS.

³⁷ Ce seuil de 92% ne représentait pas un bon pronostic pour les experts. Ils s'attendraient à une SSMD à cinq ans au moins supérieure à 96 %.

³⁸ Borne inférieure de l'intervalle de confiance 95% proche ou franchissant la valeur de HR= 0,4-0,5 dans certaines analyses rapportées.

³⁹ MINDACT : SSMD à cinq ans de 93,9 % contre TAILORx : 98 à 99 % de patientes sans métastase à distance à cinq ans.

⁴⁰ Sestak I et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer *JAMA Oncol* : 4 ; 2018.

⁴¹ Comparaison de leur valeur pronostique en analyse univariée basée sur des indices peu pertinents cliniquement (HR, rapport de vraisemblance...).

2- Éléments de réflexion et perspectives d'utilisation des SG en contexte français.

a) Clarification de la zone d'incertitude décisionnelle majeure : « zone grise » du clinicien

Les experts rappellent que les SG doivent être utilisées **au cas par cas** après avoir établi en RCP la **gravité de la pathologie** et le **niveau de certitude** vis-à-vis d'une décision de CT adjuvante **basée** sur les référentiels existants avec les CCP traditionnels.

Si selon les experts les référentiels de bonne pratique sont harmonieux pour la prise de décision thérapeutique dans la grande majorité des cancers du sein, ils ne sont globalement pas suffisamment clairs pour définir la véritable « zone grise » du clinicien.

En prenant exclusivement des CCP traditionnels recommandés en France pour la prise de décision⁴², l'ajout d'une SG en supplément (*add-on*) en cas de situation d'incertitude décisionnelle majeure pourrait cibler prioritairement un **profil plus étroit, plus homogène et mieux défini de patientes**⁴³ (proche du profil établi dans les dernières recommandations françaises de Saint Paul de Vence en 2015).

Il s'agit de patientes à profil particulièrement homogène mais **sans indication clairement établie de CT adjuvante** (peu ou pas de critères de mauvais pronostic) possédant l'ensemble des caractéristiques suivantes :

- tumeurs pT1c-pT2⁴⁴ ;
- RO + ; HER2 - ;
- avec ou sans récepteurs à la progestérone (RP) ;
- de type histologique canalaire⁴⁵ ;
- pN0-N1mic⁴⁶ ;
- de grade 2.

Il est important de noter que d'autres situations d'incertitude peuvent être perçues par certains cliniciens et représentent les cas de **discordance clinico-histo-pronostique majeure** (par ex : petite tumeur pT1b mais de grade 3 ou pN1a). Ces situations cliniques élargissent considérablement la population cible et complexifient d'autant la prise de décision. Dans ce cas, **la décision finale de CT adjuvante ne devrait pas, en l'état actuel des connaissances, s'appuyer sur le résultat d'une SG pris isolément** (résultat qui peut être corrélé à certains CCP et peu à d'autres) mais plutôt sur l'élaboration d'un **modèle de prédiction du risque (ou algorithme décisionnel) performant** incluant l'ensemble des CCP pertinents, éventuellement accompagnés d'une SG qui aura montré sa valeur ajoutée (Cf. Q2 du rapport provisoire). **Aucun modèle correspondant à cette définition n'a pour le moment été identifié ni dans la littérature de haut niveau de preuve (études d'impact clinique) ni par le GT**

⁴² Le Ki67 et l'index mitotique ont une valeur informative mais ne sont pas des marqueurs actuellement recommandés en France pour la prise de décision thérapeutique {, 2015 #124}

⁴³ Environ 5 à 10 % de patientes éligibles

⁴⁴ Les petites tumeurs lumorales infracentimétriques pT1b (pN0 grade 2) sortent du cadre de l'évaluation ciblée par la HAS (population source) notamment en raison de leur très bon pronostic global (cf. risque de récurrence de la population TAILORx à bas risque clinique) et de la nécessité de connaître le devenir clinique des formes potentiellement plus agressives mais minoritaires (cf. CR de la première réunion)

⁴⁵ Ce type histologique agressif qui est très majoritaire a été clairement privilégié par le GT pour la réalisation d'une SG (Cf. CR première réunion). D'autres types histologiques apparaissent de trop bon pronostic pour les experts (par ex : tubuleux) ou n'ont pas fait l'objet d'un consensus faute de données spécifiques identifiées dans la littérature de haut niveau de preuve (étude d'impact clinique comparative) retenue par la HAS (par ex : lobulaire).

⁴⁶ Au vu des études retenues par la HAS pour cette évaluation (absence de données probantes chez les N+) et dans l'attente des deux grands essais randomisés en population N+ (en particulier avec macrométastase) tels que RXPONDER (Oncotype Dx) et OPTIMA (Prosigna). Néanmoins, un envahissement ganglionnaire micrométastatique (pN1mic) n'est pas un critère majeur de mauvais pronostic pour le GT et modifierait peu le pronostic d'une patiente par rapport à un statut sans envahissement ganglionnaire pN0.

b) Autres propositions et perspectives faites par le GT

Sur la question relative à la désescalade thérapeutique, les experts du GT considèrent qu'une décision de CT adjuvante basée exclusivement sur une SG comme dans MINDACT n'est pas souhaitable en pratique courante française.

Propositions des experts visant à mieux encadrer l'utilisation des SG en RCP :

- **non indication** en cas de tumeurs « non lumineales » (HER2+, RO-, triple négatifs) et chez les patientes à tumeur RO+/HER2- à bas risque clinique (ces patientes à tumeur de bon pronostic sont aujourd'hui assez fréquentes en raison du programme français de dépistage précoce du cancer du sein) ;
- **contre-indication** chez les patientes à tumeur RO+/HER2- à risque clinique élevé⁴⁷ et à risque intermédiaire avec des critères de mauvais pronostic. Une désescalade thérapeutique chez les patientes qui ont une indication clairement établie de CT adjuvante basée sur les CCP traditionnels n'est pas validée à ce jour (absence de donnée clinique de haut niveau de preuve).

Selon les experts, les SG de « 1^{ère} génération » MammaPrint et Oncotype Dx sont aujourd'hui arrivées à maturité. Chacune a pu être évaluée par un essai de financement public⁴⁸, de grande taille et avec randomisation dans une indication différente⁴⁹. Pour la population de l'évaluation de la HAS, ces études ne sont pas concluantes pour établir une amélioration de l'état de santé chez les patientes pour lesquelles on a utilisé la SG en comparaison de celles pour lesquelles on a utilisé les CCP exclusivement.

Les autres SG dites de « 2^{ème} génération » sont aujourd'hui à un niveau de maturation moins avancé. L'étude OPTIMA est en cours pour Prosigna portant sur 4 500 patientes avec un financement public britannique⁵⁰ en confirmation d'une désescalade chez les patientes à risque intermédiaire ou élevé. En revanche, aucune étude d'impact clinique n'est actuellement en cours de réalisation pour Endopredict.

Les experts estiment que pour pouvoir se prononcer sur l'utilité clinique de ces deux autres SG afin d'établir leur valeur ajoutée, des **études prospectives adaptées et bien conduites** sont requises.

Cela passerait **d'abord** par :

- le recueil et l'analyse préalables des données RIHN⁵¹ pour Endopredict et Prosigna au cours de l'année 2016, 2017 et 2018 (notamment pour les tumeurs non canalaire déjà explorées par une SG et aussi pour les autres types de tumeurs en fonction du grade histo-pronostique (voire du Ki67) estimé en anatomopathologie) ;
- le renforcement de la qualité et du niveau d'exigence quant à la collecte de données des centres demandeurs et effecteurs soutenue par le RIHN (recueil exhaustif des CCP et des données cliniques de suivi) ;
- la focalisation prioritaire du recrutement dans le cadre du RIHN vers une population à tumeur lumineale homogène et en pleine « zone grise » comme définie plus haut (i.e. canalaire, pN0-N1mic, grade 2) ;
- le recueil exhaustif de données prospectives **dans la population à spectre resserrée appartenant à cette « zone grise » homogène** permettant de mieux apprécier :
 - les concordances décisionnelles entre SG, le cas échéant ;
 - les forces de corrélation et d'indépendance statistique vis à vis des CCP traditionnels ;

⁴⁷ De stade précoce ou localement avancée

⁴⁸ TAILORx : National Cancer Institut (NCI), MINDACT : Commission européenne et EORTC

⁴⁹ MINDACT - outil décisionnel exclusif pour tous profils moléculaires de tumeur ; TAILORx - confirmation d'une désescalade chez des patientes à risque faible et intermédiaire de récurrence

⁵⁰ National Health Service (NHS)

⁵¹ Référentiel des actes innovants hors nomenclatures

- ▶ le spectre de dispersion ou de distribution des valeurs observées de scores génomiques « en vie réelle » (sans tenir compte du seuil décisionnel proposé par le fabricant)

De l'avis général du GT, l'absence de données permettant de quantifier concrètement leur valeur ajoutée ferait des SG à terme un simple examen de confort coûteux sans intérêt clinique démontré pour la patiente.

c) Identification d'un risque de désinformation des patientes et des professionnels concernés

Un dernier point mais néanmoins majeur souligné par les experts est la dérive informationnelle véhiculée dans la presse grand public de l'inutilité - voire du caractère délétère - de 70 % des CT adjuvantes actuellement prescrites en France dans le cancer du sein depuis l'accès rendu possible aux SG.

Le GT a tenu à souligner la dangerosité de ces messages diffusés à destination de la population et des professionnels concernés. Ces messages ne prennent pas en compte l'apport établi des CCP traditionnels et du contexte de soins français en particulier avec la mise en place depuis de nombreuses années d'une réunion de concertation pluridisciplinaire (RCP) et du contrôle qualité des laboratoires d'analyses médicales.

Selon les experts, **en dehors du cadre d'une RCP et en l'absence d'une information loyale claire et appropriée avec accord de la patiente**, aucun professionnel de santé référent ou non de la patiente ne devrait prescrire une SG et l'en informer de ses résultats.

Annexe 12. Retour *in extenso* des parties prenantes

Réponse sous format libre de la Société Française de Sénologie et de Pathologie Mammaire (SFSPM)

Ce rapport complexe a fait l'objet de commentaires détaillés de la part de quelques membres de notre société particulièrement impliqués dans ce domaine.

En préambule, il faudrait préciser que le contexte nord-américain et la situation française sont différents pour de multiples raisons, ce qui peut conduire à une utilisation différente de ces signatures par les cliniciens (1)

REMARQUES DIVERSES :

Certains termes ne sont pas clairement définis : « tumeur luminale » dont la définition ne semble apparaître que page 72, et « hormonodépendance », qu'il faudrait plutôt remplacer par « analyse immunohistochimique confirmant des RE et/ou RP positifs ».

Par ailleurs, les niveaux d'expression des RH et des Her2 ne permettent pas d'évaluer « la réponse » à l'hormonothérapie et aux thérapies ciblées anti-Her2 mais le « bénéfique » de ces traitements en terme de SSR et SG.

De façon étonnante, parmi les référentiels internationaux (1.5 page 17) les deux rapports majeurs de l'ASCO ne sont pas cités (HARRIS LN, JCO 2016 et KROP J, JCO 2017), de même que les guidelines du NCCN et de l'ESMO.

D'autre part, le critère « âge jeune » (< 35/40 ans) n'est plus de façon isolé un élément indiquant obligatoirement une chimiothérapie, et a été retiré des recommandations internationales.

Pour la question N°2 le libellé devrait se focaliser sur le pronostic des signatures et pas sur la prédiction. De façon plus générale, il faudrait donc préciser que les signatures ont été développées sur leur valeur pronostique (cf. essais prospectifs) et qu'aucune ne permet la prédiction précise du bénéfique de la chimiothérapie.

Par ailleurs, il faut souligner que les grands essais (MINDACT, TAILOR-X) ont été conçus il y a plus de 10 ans, au début de l'utilisation des signatures moléculaires, avec des populations probablement pas assez ciblées sur la vraie « zone grise » d'incertitude thérapeutique.

Il faudrait enfin discuter des recommandations d'autres pays, notamment du NICE (www.nice.org.uk/guidance) qui utilisent le test PREDICT (version 2.0) et ONCOTYPE DX™ ; ce dernier test pour les patientes à risque « intermédiaire » pN0 RO+ Her2-.

(1) : exemple : contrôle de qualité en anatomopathologie

Réponses de Société Française de Chirurgie Oncologique (SFCO)



Développer la qualité dans le champ
sanitaire, social et médico-social

**RELECTURE DU DOCUMENT PROVISOIRE
ET CONFIDENTIEL INTITULÉ**

**« *Utilité des signatures génomiques dans le cancer du
sein de stade précoce* »**

Septembre 2018

Nous vous remercions d'avoir accepté de participer à la relecture de cette évaluation en répondant à ce questionnaire.

Nous vous saurions gré de bien faire apparaître le point de vue de votre organisme et autres éclairages éventuels dans ce document.

Vos commentaires seront intégralement reproduits dans le rapport définitif d'évaluation que la HAS rendra public à l'issue de son processus d'évaluation. Jusqu'à cette échéance, l'argumentaire qui vous a été transmis demeure par conséquent strictement confidentiel.

Nos contraintes calendaires d'évaluation nécessitent que vous nous retourniez votre réponse par voie électronique avant le **28/09/2018** (has.seap.secretariat@has-sante.fr).

LISIBILITÉ ET CLARTÉ

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

- L1** Réponse : synthèse page 44 manque de clarté. On comprend bien les contre-indications aux SG. Par contre on ne comprend pas très bien si la population en incertitude décisionnelle majeure est éligible à la réalisation de SG. Quel test ?

POINT DE VUE DE VOTRE ORGANISME

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

- P1** Réponse : Clarifier la « synthèse » page 44.

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

- P2** Réponse :

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

- R1** Réponse : Souligner l'importance d'homogénéiser les pratiques de prescription de CT en France (rapport de l'IRDES n° 221 - Janvier 2017). Les SG sont un levier à la pertinence des soins sous réserve que l'on respecte leurs indications.
Les référentiels de bonnes pratiques sont en effet indispensables, mais en pratique il n'y en a pas au niveau national et encore moins régulièrement actualisés. Suivons ceux qui existent au niveau européen voire du NCCN.

Réponses de l'association de patients « La Ligue nationale contre le Cancer »

LISIBILITÉ ET CLARTÉ

L1

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Réponse :

Le document fourni nous paraît très clair et bien structuré.

POINT DE VUE DE VOTRE ORGANISME

P1

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Réponse :

La Ligue Nationale Contre le Cancer (LNCC) a beaucoup apprécié la grande rigueur de l'évaluation fournie par la HAS.

L'évaluation de la HAS souligne le manque d'études permettant de porter des conclusions fiables sur les 3 questions cliniques (Q1-Q3) qui sont soulevées ; elle indique en creux les paramètres qu'il faudrait prendre en compte pour construire des études à même de mener à des conclusions fiables.

Dans cet esprit, la LNCC trouverait très utile que la HAS adjoigne à son évaluation une annexe décrivant le design idéal du ou des essai(s) permettant en théorie de porter des conclusions fiables sur ces 3 questions (Q1-Q3). Ce complément serait très appréciable pour mesurer la distance restant à parcourir pour obtenir des conclusions sur ces questions importantes. Il aiderait aussi à orienter plus efficacement les études futures sur ces questions.

P2

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Réponse :

L'évaluation de la HAS aboutit à une position qui peut être qualifiée de conservative relativement à l'utilisation des Signatures Génomiques commerciales. Un effet potentiel de cette position peut-être de décrédibiliser l'apport possible de la génomique en oncologie, à rebours des espoirs soulevés par le plan France Médecine Génomique 2025. La LNCC participe depuis longtemps au financement de recherches en médecine génomique. Un effet potentiel du rapport de la HAS pourrait être de décourager le financement de ce type de recherches par la LNCC, faute d'adhésion de ses membres quant à l'intérêt de ces approches.

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

Réponse :

Remarque 1 :

La LNCC comprend l'intérêt pratique de la question 4 (Q4). La LNCC trouve qu'une autre question, connexe, aurait largement mérité de retenir l'attention du groupe de travail : « Peut-on, pour chacune des questions Q1-Q3, hiérarchiser les signatures existantes quant à leur capacité à répondre à l'objectif proposé ? ». En effet, si l'on pouvait établir que certaines signatures commerciales sont clairement moins ou plus performantes que les autres (dans le contexte des questions Q1-Q3), il apparaîtrait très important de le faire largement savoir. La LNCC peut imaginer que la HAS puisse être frileuse à l'idée de soulever cette question, du fait des batailles juridiques avec les industriels que pourraient entraîner les conclusions. Cependant, il apparaît à la LNCC que l'intérêt prioritaire à défendre est celui de l'information des patient(e)s, et non la tranquillité à court terme des institutions d'évaluation quelles qu'elles soient.

R1

Remarque 2 :

La LNCC apprécierait aussi beaucoup que cette évaluation intègre des informations sur le degré de reproductibilité de la décision thérapeutique :

- *(i) dans le processus décisionnel actuel, c'est-à-dire hors utilisation de signature génomique,*
 - *(ii) dans le(s) processus décisionnel(s) intégrant une signature génomique,*
- ... en tenant compte chaque fois de toute la chaîne d'actions nécessaires à la prise de décision. En effet, il se pourrait que les deux processus décisionnels (avec versus sans signature génomique) aient une efficacité moyenne comparable, mais un degré de reproductibilité globale différent. Cette information nous apparaît très importante dans le cadre de cette évaluation.*

Réponses du Conseil National Professionnel des Pathologistes (CNPP)

LISIBILITÉ ET CLARTÉ

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Le Collège National Professionnel des Pathologistes (CNPath) a lu avec intérêt le rapport provisoire du groupe de travail réuni par l'HAS à propos de l'utilité clinique des signatures génomiques dans le cancer du sein de stade précoce. Il en a apprécié notamment la qualité générale, l'état des lieux sur les principaux rapports d'évaluation étrangers (chapitre 1.5) et la présentation détaillée et argumentée de l'analyse critique des données sélectionnées dans la littérature par le groupe de travail. Ce rapport provisoire est lisible et clair. Ses conclusions, selon lesquelles il n'y a pas d'utilité clinique actuellement démontrée pour les signatures moléculaires existantes dans le cancer du sein, sont données sans ambiguïté. Il n'existe cependant pas de conclusion claire sur les perspectives à court et moyen terme à envisager pour ces analyses : doivent-elles être suspendues définitivement ou en attendant des données supplémentaires ? doivent-elles être poursuivies de manière encore plus encadrée ?

Plusieurs remarques peuvent être faites.

A. Remarques majeures

- la méthodologie de travail est clairement explicitée ; cependant, dans la mesure où deux réunions physiques seulement ont été tenues, la part de travail réalisée en dehors des réunions physiques mériterait d'être mieux détaillée, dans son organisation et sa méthode

L1

- certains mots clés méritent une définition plus précise : c'est le cas du terme « preuve directe » souvent employé, ainsi que de celui de « haut niveau de preuve », employé notamment page 33.

- il est fait allusion au « contexte français » en fonction duquel les résultats des études multicentriques internationales (où la France a souvent participé de manière significative) doivent être mises en perspective ; il serait intéressant de mieux définir les particularités de ce contexte par rapport à celui d'autres pays (référentiels, organisation des soins, financement ?) et la façon dont elles amènent à pondérer les résultats de la littérature

- Le contenu des signatures moléculaires qui constituent l'objet même du rapport, leur genèse, les algorithmes les définissant ne sont ni détaillés, ni commentés.

- Les termes de « profil moléculaire », voire de « statut moléculaire » (page 61), sont souvent employés pour désigner le phénotype immunohistochimique des cancers du sein, qui constitue une base essentielle du raisonnement clinique ; sans être impropres (les protéines mises en évidence par immunohistochimie sont bien des molécules !), ces termes peuvent cependant prêter à confusion avec les profils ou les statuts « moléculaires » définis par des analyses génétiques. De même, il existe des confusions sur la définition du phénotype luminal tout au long du rapport (la définition retenue est-elle transcriptomique ou immunohistochimique ?).

- Certaines informations données dans les annexes 7 et 8 ne sont pas reprises dans le texte principal

B. Remarques mineures

- le terme d'invasion lymphovasculaire « extensive » est plusieurs fois utilisé parmi les facteurs histologiques de mauvais pronostic, mais sans qu'une définition claire (ou une référence précise) ne soit donnée
- la Figure 2 pourrait être améliorée pour gagner en clarté et mettre mieux en valeur les questions posées par le groupe de travail
- il existe d'assez nombreuses coquilles typographiques, notamment dans le chapitre 1, qu'une relecture attentive permettra d'éliminer

POINT DE VUE DE VOTRE ORGANISME

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Le CNPath se félicite de voir souligné le rôle essentiel joué par les critères issus de l'examen anatomopathologique dans l'évaluation pronostique du cancer du sein. Il apprécie l'analyse critique qui est faite des données de la littérature sélectionnées par le groupe de travail à propos de la validité clinique et l'utilité éventuelle des signatures moléculaires actuellement disponibles.

Plusieurs remarques peuvent cependant être faites.

- Les signatures moléculaires ont été initialement conçues pour une évaluation pronostique, et non pas pour une évaluation prédictive de la réponse aux traitements, notamment aux chimiothérapies. Les deux aspects devraient être mieux distingués dans l'analyse des études disponibles et dans les conclusions qui en sont tirées par le groupe de travail

- Les questions posées par le groupe de travail sont pertinentes, et notamment la question 2. Cependant, le groupe de travail a été amené à rechercher les réponses à ces questions dans des études cliniques qui n'avaient pas nécessairement les mêmes objectifs. De même, le groupe de travail ne semble pas tenir compte de ce que certaines études (comme MINDACT ou TAILORx) ont été conçues il y a plus d'une décennie, au moment de l'émergence des signatures moléculaires, avec une population volontairement large. Il est difficile d'en attendre des réponses aux questions précises qui se posent aujourd'hui, avec le recul et l'expérience acquise. Ces points doivent être pris en compte et discutés dans le rapport.

- le groupe de travail base l'essentiel de ses conclusions sur un nombre limité de publications internationales et notamment, sur deux études prospectives ; il n'a pas retenu certaines informations, comme les données des registres prospectifs tels que Raster pour Mammaprint® (Int J Cancer 2013;133:929-36) ainsi que les données des bases SEER (Breast Cancer Res Treat 2017;163:303-31), CLALIT (Breast Cancer 2017;3 :art32) et DCBG (J Clin Oncol 2018; 36 :735-740) ; ces données incluent des informations cliniques et pronostiques dans des conditions « actuelles » qui pourraient être utilement comparées avec les données des essais cliniques ou des essais prospectifs évalués de façon rétrospective.

- le rapport provisoire rappelle les résultats de plusieurs groupes de travail étrangers (chapitre 1.5) ayant conclu, comme lui, à l'absence de « preuve directe » en faveur de l'utilité clinique des signatures moléculaires ; cependant, d'autres groupes de travail ont récemment rendu des conclusions différentes : analyses européennes NICE (<https://www.nice.org.uk/guidance/indevelopment/gid-dg10015/documents>), groupe de travail allemand (<https://www.iqwig.de/en/press/press-releases/biomarker-tests-in-breast-cancer-new-study-data-indicate-advantage-for->

certain-patients.10059.html), sociétés espagnoles d'oncologie médicale et de pathologie (Clin Transl Oncol (2018) 20:815–826 ; <https://doi.org/10.1007/s12094-017-1800-5>); ces conclusions différentes méritent d'être discutées dans le rapport, ce qui renforcera les positions prises par le groupe de travail

- le groupe de travail a fait volontairement le choix de limiter ses référentiels aux recommandations internationales publiées avant l'entrée en force des signatures moléculaires (avant 2009 en substance) ; il explicite certaines raisons de sa position dans le rapport provisoire; néanmoins, tous les praticiens, y compris les pathologistes, ont accès à des recommandations récentes (comme ASCO, St Gallen) dont les propositions concernant l'usage des signatures moléculaires sont parfois en contradiction apparente avec les conclusions du rapport provisoire. De plus, les pathologistes sont familiers avec la classification TNM (AJCC/UICC), internationalement reconnue et appliquée, dont la dernière édition tient compte de l'existence des signatures moléculaires. Ce point devrait être explicité dans le rapport définitif, où les éventuelles discordances pourraient être justifiées.

Enfin, le CNPath constate le ton particulièrement critique utilisé non seulement dans l'analyse de certaines études cliniques (notamment, page 37) mais aussi dans l'évaluation générale des conditions de leur conception et de leur publication (chapitre 3.3.3, page 42). Il s'interroge sur l'utilisation polémique qui pourrait en être faite et sur les conséquences qu'il pourrait avoir sur la réception des conclusions du rapport.

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Le sujet des signatures moléculaires dans le cancer du sein concerne les pathologistes à plusieurs niveaux :

1- la contribution à la sélection des patientes pour lesquelles le test est réalisé : comme le rappelle le rapport provisoire, cette sélection est notamment basée sur des paramètres anatomopathologiques qui sont des facteurs pronostiques clés, dont, pour la plupart, la valeur et l'utilité cliniques sont bien démontrées

A cet égard, plusieurs paragraphes du rapport provisoire sont consacrés à l'index Ki-67 pour rappeler les limites de sa valeur pronostique et, probablement pour faire le parallèle avec les signatures moléculaires, l'absence d'utilité clinique clairement démontrée ; sans remettre en cause les arguments présentés, avec lesquels il est globalement en accord (même si ce paramètre est attendu par la plupart des cliniciens), le CNpath s'interroge sur la nécessité de développer plusieurs fois ce point, qui n'est pas dans les objectifs affichés du groupe de travail

2- la réalisation pratique du test: c'est en effet le pathologiste qui sélectionne l'échantillon tissulaire à partir duquel le test moléculaire est réalisé et l'adresse aux plateformes pratiquant l'analyse ; au-delà, certains pathologistes participent directement ou indirectement à la réalisation même du test et/ou à son interprétation

Le CNPath constate que, dans le rapport provisoire, les critères requis pour la sélection de l'échantillon tissulaire à partir duquel les tests moléculaires sont réalisés ne sont pas discutés. Plusieurs paramètres, comme les conditions pré-analytiques, la cellularité tumorale, la représentativité de l'hétérogénéité tumorale, s'il y a lieu, peuvent éventuellement constituer des facteurs confondants pour la réalisation des tests moléculaires et l'interprétation de leurs résultats. Ce point pourrait être souligné, d'autant qu'il est explicitement abordé dans l'autosaisine de l'HAS. L'évaluation de la qualité de ces pratiques doit aussi être discutée voire requise.

3- l'implication dans le parcours de soin : c'est en effet au pathologiste que s'adresse

P2

l'oncologue lorsqu'il souhaite la réalisation d'un test génomique et le pathologiste participe aux RCP dans lesquelles la décision de recourir à un test moléculaire est prise

A cet égard, il est important que le pathologiste dispose des mêmes sources d'information que les autres praticiens. La pression pour réaliser ces tests est forte mais bien maîtrisée, notamment depuis leur inscription au RIHN. Tout « retour en arrière » sur leurs indications et sur les modalités de leur prise en charge financière devra être largement diffusé et expliqué à la communauté médicale, incluant les pathologistes, qui sont la cible principale des demandes et des pressions exercées pour la réalisation des tests. Il est important de souligner ici que, dans le cadre du RIHN, les demandes de test moléculaire sont actuellement encadrées par la RCP pour une population très ciblée (« population en incertitude décisionnelle »), avec des critères de prescription très majoritairement respectés (cf Tableau 2, page 16).

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

R1

Le CNPath regrette que l'inscription dérogatoire des signatures moléculaires au RIHN ne permette pas de recueillir les informations nécessaires à une évaluation de l'utilité clinique de ces tests, contrairement aux objectifs mêmes de cette procédure. Il constate également que l'autosaisine de l'HAS a été décidée avant la fin de la période de 3 ans officiellement prévue pour l'évaluation d'une analyse inscrite au RIHN. Les raisons de cette autosaisine sont expliquées dans le document correspondant disponible sur le site de l'HAS. Dans ce contexte, le CNPath réitère son attachement à l'évaluation objective, factuelle et scientifique de tout nouveau biomarqueur, sans en oublier la dimension médico-économique, et souhaite que le rapport définitif fournisse à tous les praticiens impliqués, dont les pathologistes, les réponses claires et argumentées dont ils ont besoin, mises en perspectives dans le contexte international. De telles analyses pourraient être concertées avec les autres agences européennes, qui ont déjà rendu leurs rapports en ne se limitant pas aux seules études prospectives.

Réponses du Conseil National Professionnel de Biologie Médicale (CNPBM)

Concernant les remarques mineures signalées par les experts du CNP de Biologie Médicale :

- Les figures sont floues à l'impression
- La figure 1 n'indique pas le grade et l'histologie de type « canalaire » comme cela est indiqué dans le reste du document.
- Tableau 2: clarifier « test centralisé » : hors de France ?
- Page 31: clarifier la figure 3: on part d'une sélection de 11 publications, on en élimine 7, il devrait en rester 4, or 5 publications sont retenues ?
- Figure 3: indiquer clairement les références des 5 publications retenues.
- Une faute de Frappe: p32 paragraphe sélection bibliographique: je pense qu'il faut lire "des traitements HT ou CT/HT" et non "des traitements CT ou CT/HT"
- Une faute de Français: p66 paragraphe 6: "malgré qu'ils puissent"

LISIBILITÉ ET CLARTÉ

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Réponse:

Expert 1 : *Le rapport est détaillé et volumineux et les conclusions les plus claires figurent à l'annexe 8 à la fin du compte rendu de la réunion du GT du 02 juillet 2018. Ces éléments pourraient être repris et mis en exergue dans une conclusion du rapport lui-même.*

L1

Expert 2 : *Il serait souhaitable de disposer d'un tableau de synthèse comparant les caractéristiques des 4 types de signatures génomiques avec les paramètres suivants :*

- Nombre de transcrits testés
- Méthodologie : RT-PCR, hybridation...
- Automatisation, standardisation, praticabilité technique
- Possibilité de réalisation de contrôles de qualité par les biologistes médicaux
- 1^{er} ou 2^{ème} génération
- Ajouts ou non dans l'algorithme de ces 4 SG de divers CCPs : taille macroscopique, envahissement ganglionnaire...
- Réalisation ou non dans les laboratoires en France.

POINT DE VUE DE VOTRE ORGANISME

P1

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Réponse :

Expert 1 : *Si le questionnement est bien posé et précis, l'argumentation sur les*

2 seules études actuellement publiées relève d'une technicité importante et s'adresse à des méthodologistes confirmés.

Expert 2 : Cette évaluation est fragilisée par le fait qu'elle repose sur

- uniquement sur 5 publications retenues
- sur des études financées par des industriels (Ex : Genomic Health)
- sans aucune publication comparant simultanément les 4 SGs
- sans aucune publication concernant le test Endopredict.

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Réponse :

Expert 1 : Les conclusions de cette évaluation sont clairement en faveur d'une absence actuelle de justification médicale de la prescription des signatures génomiques commerciales actuellement disponibles dans le cancer du sein. Celles-ci relèvent donc encore de la recherche clinique et la cohérence voudrait que leur financement n'incombent pas aux budgets attribués au RIHN mais dépendent de programmes spécifiques relevant par exemple du PHRC.

P2

Expert 2 : Il aurait été souhaitable de :

- Faire des recommandations concernant le choix des tests à utiliser (des tests centralisés hors de France ne devraient pas pouvoir être inscrits au Référentiel des actes innovants hors nomenclatures (RIHN))
- Proposer un arbre décisionnel clair pour les indications de ces tests SG en fonction de l'ensemble des CCPs (incluant le Ki67) - Ex : propose-t-on une GS pour un **grade I : T2N0N1** ?
- Proposer des indications claires, non ambiguës et **très restrictives** pour l'inscription au Référentiel des actes innovants hors nomenclatures (RIHN). Il ne faut pas que ces tests très coûteux altèrent fortement l'enveloppe financière RIHN.

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

Réponse :

Expert 1 : L'évaluation réalisée ici aborde la question de l'intérêt clinique des signatures génomiques dans le cancer du sein d'un point de vue strictement empirique ce qui est parfaitement légitime pour la question posée (demande de démonstration d'un gain médical par l'expérience directe dans une situation précise). On peut cependant regretter qu'aucune tentative d'explication biologique ne soit apportée et qu'à aucun moment ne soient données des précisions sur la nature des signatures recherchées ni sur leur mode d'obtention. Cette approche aurait peut-être permis de préciser les phénomènes biologiques sous-jacents analysés par ces outils et de comprendre en quoi l'information obtenue était finalement redondante avec les éléments phénotypiques utilisés en pratique médicale courante pour apprécier le pronostic et les indications thérapeutiques dans le cancer du sein. Cette dimension de tentative de compréhension de la biologie sous-jacente aurait été bienvenue.

R1

Expert 2 : Cette évaluation réalisée par la HAS semble très préliminaire (peu de publications fiables et ne concernant que certaines des SGs - les 5 publications retenues sont très complexes, difficiles à lire et à interpréter). Des recommandations fiables ne peuvent donc pas résulter de cette évaluation. Il semble indispensable de réaliser une large étude clinique prospective (sur le modèle de l'essai Optigen qui n'a pas encore commencé), qui serait piloté par un organisme indépendant (Ex : INCa) sans aide financière par des industriels (Ex : Genomic Health ; Myriad Genetics...) comparant les 4 tests SGs + le test IHC4 dans une population restreinte et parfaitement définie.

Réponses du Groupe Francophone de Cytogénomique Oncologique (GFCO)

LISIBILITÉ ET CLARTÉ

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Réponse :

Le contexte est posé en préambule et les 4 questions, sujet de l'évaluation, sont clairement exposées. L'étude méthodologique est exhaustive, critique et claire.

Cependant, l'analyse peut paraître limitée compte-tenu du faible nombre de publications retenues avec la méthodologie de travail choisie, d'exclusion des études rétrospectives.

La population d'indication des tests est bien discutée mais la limitation d'usage des SG à cette dernière doit être plus clairement indiquée. Dans la synthèse de la position du GT (chapitre 3.3.5), les indications des SG (à savoir les cancers du sein luminaux à risque intermédiaire) se déduisent plutôt des non-indications énoncées. Ce point devrait apparaître en clair. Il est également important d'indiquer plus clairement la contre-indication à l'utilisation des SG dans les tumeurs non-luminales ou les tumeurs ayant une indication de prise en charge claire sur la base des facteurs CCP.

Les conclusions des 4 questions posées sont un manque de preuves directes de l'utilité clinique des SG. Elles sont sans ambiguïté, particulièrement pour les questions 1 et 3.

On peut cependant noter que certains points importants, clairement discutés en annexe 8, ne sont pas repris dans le rapport. Notamment, de l'avis des experts, ces signatures possèdent bien une valeur pronostique intrinsèque. A noter que la première phrase du paragraphe 3.3.5 est en contradiction avec cette position du GT et non conforme à ce qui est rapporté dans l'encadré de l'annexe 8 - p74. Egalement, le GT conclu à l'absence de démonstration de la valeur prédictive d'efficacité de la chimiothérapie adjuvante pour les signatures de 1^{ère} génération (annexe 8). Enfin, dans les rapports du GT d'experts, différents niveaux de preuve retenus et mentionnés entre les quatre SG, éléments qui ne sont pas repris pour faire apparaître une prise de position entre les différentes SG.

Le GT recommande clairement de recentrer les prescriptions des SG sur les RCP pour harmoniser les prescriptions. Bien que ce rapport de la HAS n'a pas vocation à émettre des recommandations professionnelles, certaines des conclusions du rapport pourraient conduire à une utilisation plus encadrée de ces tests. Ceci doit être clairement mentionné.

Enfin, dans l'état actuel de ce document, il manque la conclusion quant à la prise en charge future des SG par l'Assurance Maladie. Egalement, en cas d'absence d'indication de remboursement par l'Assurance Maladie, aucune autre option de financement de ces tests n'est envisagée, alors même qu'elles font maintenant partie de la prise en charge des cancers du sein en France. On peut également regretter que cette évaluation n'ait pas eu lieu avant l'inscription des SG sur la liste des RIHN car elle aurait évité la diffusion des SG si elles sont jugées sans bénéfice ajouté et des dépenses de Santé conséquentes.

En l'absence de conclusion sur la prise en charge des SG, le GFCO ne peut donc pas se prononcer sur la clarté de la conclusion définitive du rapport.

POINT DE VUE DE VOTRE ORGANISME

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Réponse :

Ce rapport fait clairement ressortir qu'il existe bien une sous population de patientes pour lesquelles, malgré les outils validés actuels, le clinicien se trouve dans une situation d'incertitude décisionnelle majeure. C'est dans ce contexte restreint qu'il faut évaluer la pertinence des SG et produire des données consolidées, notamment pour les SG de 2ème génération.

Ces SG font maintenant partie intégrante de la prise en charge des cancers du sein en France. Ainsi, la position va être polémique sur les points suivants : la conclusion sur la base de 5 publications, l'exclusion de nombreux articles publiés, la remise en cause de conclusions d'essais internationaux publiés dans des revues internationales de qualité reconnue et la non prise en considération de certains référentiels internationaux à forte portée et incluant les SG (ASCO, NCCN, St Gallen, AJCC/UICC, etc...). A ce titre, il est également regrettable de ne pas pouvoir disposer de l'évaluation en cours du NICE pour une harmonisation européenne.

On peut regretter que, bien que conduite de façon rétrospective mais dans le cadre d'une étude prospective, la seule évaluation comparative des SG entre elles sur un même échantillon n'ait pas été prise en compte pour aider à répondre à la question 4 (publication de Sestak dans JAMA Oncol 2018, citée comme étant une étude de référence par le GT, en annexe 8).

Par ailleurs, on peut noter l'absence de prise en compte des données en vie réelle au niveau national, remontées de façon académique pour certaines SG. Ce mécanisme de remontée d'information du RIHN n'a clairement pas fonctionné de manière homogène entre les SG alors que l'accès dérogatoire au financement n'était pas différencié. Il aurait dû aider à l'évaluation.

Le GFCO considère qu'en l'état, le rapport n'émet pas de conclusions définitives quant à la question posée, à savoir l'utilité et donc in-fine l'éligibilité pour un remboursement de ces tests (tous ou certains et dans quel contexte, pronostic ou prédictif) par l'Assurance Maladie.

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Réponse :

Au regard de l'implantation de ces tests dans la prise en charge des cancers du sein précoces en France, tout avis de la HAS aura un impact important. Même si cela n'est pas l'objectif de ce rapport, et quel que soit le mode de financement retenu, toute absence de recommandations claires concernant la pratique en routine de ces tests, risque de conduire à une grande hétérogénéité des pratiques et la réalisation de tests indépendamment du recueil de données. Ceci n'est plus souhaitable au regard des preuves qui sont encore à apporter. Le GFCO, en tant que société savante de génomique du cancer, joue un rôle important dans l'harmonisation des pratiques, l'assurance qualité et la formation des professionnels et souhaite que les pratiques appliquées à ces SG répondent aux exigences de la Biologie Médicale en

France.

Egalement, le GFCO attire l'attention sur le risque de forte inégalité d'accès à ces tests, largement médiatisés sans positionnement plus claire.

Dans ce contexte, le GFCO pose la question du maintien de ces SG au financement RIHN avec une définition précise de la population de patientes concernées, une prescription exclusive en RCP et l'obligation de recueil et de remontée des données nécessaires à une analyse en vraie vie (également rapporté en annexe 7 et 8 mais non repris dans le rapport lui-même). En effet, il semble opportun de limiter la prise en charge à la population ciblée et imposer le recueil de données complémentaires. Cependant, la durée des inclusions pour les études prospectives à mettre en place dépassera le cadre limité à 3 ans du RIHN. Les données ne seront analysables au mieux qu'avec une médiane de suivi de plus de 5 ans.

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

Réponse :

L'auto-saisine de l'HAS sur l'utilité des signatures génomiques dans le cancer du sein de stade précoce est un premier signal fort sur le mécanisme d'évaluation annoncé lors de la mise en place des RIHN. Il est regrettable que de nombreux marqueurs avec des niveaux de preuve non contestables ne bénéficient toujours pas d'une évaluation et d'un passage à la nomenclature (marqueurs associés aux autorisations de mise sur le marché des médicaments et marqueurs utilisés à titre diagnostique dans les cancers rares).

R1

Dans le cadre dérogatoire du RIHN, l'utilisation des SG est quasiment rentrée dans la pratique standard en France. Il sera difficile d'apporter des preuves complémentaires sur la base de nouvelles grandes études internationales randomisées sachant que peu de marqueurs disposent d'une telle littérature.

Egalement, dans le contexte de la désescalade thérapeutique (question 3), il aurait pu être posé la question de l'impact médico-économique et l'impact en termes de qualité de vie pour les patientes.

A noter

Une coquille p34 sur l'étude TAILORx : « traitement CT exclusive ou CT/HT » au lieu de HT versus CT/HT.

p44 : avis général du GT non cohérent avec le CR du GT p74. La valeur pronostique n'a pas été remise en cause par le GT.

Réponses du Société Française de Médecine Prédictive et Personnalisée (SFMP)

LISIBILITÉ ET CLARTÉ

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Réponse :

L1

Les importantes remarques de fond qui suivent, portant sur la définition des objectifs, l'analyse critique des données, l'absence de discussion génomique et de discussion des référentiels internationaux, la confusion recherche/RIHN et un désaccord avec les conclusions sont telles que nous répondrons surtout sur le fond. Le document, malgré un travail important des experts, n'est au final ni clair, ni objectif. À titre d'exemple, l'utilisation de qualificatifs tout au long du rapport n'est pas classique dans un travail d'expertise scientifique et donne une impression de subjectivité.

POINT DE VUE DE VOTRE ORGANISME

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Réponse :

La SFMP remercie la HAS de l'avoir sollicitée pour la relecture de ce rapport. Ce dernier contient un volume important d'informations et d'avis. Toutefois, la SFMP est en désaccord sur de nombreux points d'étapes d'analyse et d'interprétation des données de la littérature et par conséquent des conclusions émises, ce qui sera détaillé plus loin dans ce rapport de relecture. La SFMP formule un avis très réservé sur l'analyse et les conclusions du rapport d'évaluation. La SFMP, qui aurait souhaité se limiter à des critiques de forme et de fond constructives pour ce rapport (cf. R1), est au regret de devoir exprimer un désaccord allant au-delà et ne peut pas s'approprier ce travail.

P1

Il nous semble de plus important d'attirer l'attention de la HAS, Autorité publique indépendante agissant au nom de l'état, qu'elle s'expose à un risque de contestation non seulement de la part de la communauté médicale mais également des patientes si ce rapport est publié dans l'état actuel.

En effet, la SFMP qui a pour objectif de promouvoir les progrès scientifiques dans le domaine de la médecine prédictive et personnalisée s'interroge sur les conséquences de ce rapport qui n'est pas en ligne avec l'intérêt des malades, ni avec le rôle à venir de la médecine génomique ou de précision souligné par Madame la Ministre de la Santé, le plan France Médecine Génomique et le Conseil National de l'Ordre des Médecins.

Brièvement, les cinq experts consultés par la SFMP estiment que le groupe de travail a eu une lecture non pertinente des données disponibles, notamment en ne tenant pas compte des niveaux de preuve et des validations prospectives récentes des signatures génomiques. Il est incompréhensible d'un point de vue méthodologique

que le groupe de travail du rapport de l'HAS propose de rester uniquement à un conseil basé sur des facteurs biopathologiques, dont certains ne sont pas validés de façon prospective, en rejetant ceux de certaines signatures de niveau de preuve I évalués eux de façon prospective.

Le rapport comprend des biais d'analyse liés à des erreurs assez grossières de définition des objectifs des études et/ou des populations étudiées (cf R1).

La méconnaissance des aspects moléculaires du sujet de ce rapport est manifeste avec aucunes considérations (ni rappel, ni discussion) sur le contenu génomique des signatures et des confusions notables sur la définition du phénotype luminal tout au long du document. Il ne semble pas qu'un spécialiste en médecine génomique ait participé aux travaux.

Par ailleurs, la non considération des données récentes du plus haut niveau de preuve est non seulement en décalage avec l'état des connaissances actuelles mais aussi avec les référentiels internationaux. De façon étonnante ceux-ci ne sont d'ailleurs pas cités dans le rapport.

Contrairement aux conclusions principales, la SFMPP pense souhaitable d'intégrer rapidement en pratique les signatures génomiques, dont certaines sont aujourd'hui validées pour certains sous-groupes de patientes. Ils démontrent dans ces groupes de femmes, l'absence d'intérêt de la chimiothérapie adjuvante, permettant ainsi d'éviter des effets secondaires aux femmes et certaines dépenses inutiles. Ceci impacte donc positivement les objectifs d'optimisation des dépenses de notre politique de santé publique et permet avec la sécurité suffisante pour les patientes, et une confiance importante pour les oncologues, des décisions de désescalade thérapeutiques raisonnées.

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Réponse :

La SFMPP souhaite attirer l'attention de la HAS sur des réactions négatives possibles des associations de patientes, des sociétés savantes et de la communauté médicale, si ce rapport reste en l'état.

P2

Ne pas aller vers une ouverture à l'intégration des signatures dans la pratique, placera non seulement **la France en décalage par rapport aux recommandations internationales**, mais surtout pourra constituer une **perte de chance pour les patientes** et une réelle **difficulté d'exercice pour le médecin qui a obligation de moyens**. En outre, elle créera une **non équité d'accès aux soins**, puisque certains établissements ou patients pourront financer ces analyses et d'autres non.

Des preuves que les signatures peuvent être utiles en vue de la désescalade (c'est-à-dire éviter des chimiothérapies inutiles) sont là, il n'est pas éthique que celles-ci soient ignorées ou rejetées de façon subjective, alors que les études en regard comptent parmi les plus importantes (qualitativement et quantitativement) de l'histoire de la cancérologie moderne.

Les propos dénigrant les deux essais prospectifs randomisés majeurs, allant jusqu'à **questionner sur l'intégrité scientifique de ces études et des investigateurs ne nous paraissent pas acceptables** pour un rapport émanant d'une agence gouvernementale. Rappelons que le NCI est le sponsor de l'étude TaylorX et que l'Europe

(CEE) via le BIG est le sponsor de l'étude Mindact (avec une participation française significative pour les inclusions).

REMARQUES LIBRES

Avez-vous d'autres commentaires à formuler ?

Réponse :

Commentaire général

Les deux études prospectives randomisées majeures portant sur l'utilité de la chimiothérapie dans le cancer du sein précoce constituent des événements majeurs depuis la dernière expertise des signatures par l'INCa. **Il est incompréhensible que le GT n'ait pas pris en compte ces essais récents comme des avancées significatives pour le transfert clinique. Rappelons que ces essais ont le niveau de preuve le plus élevé, comptent parmi les plus grandes études prospectives randomisées jamais réalisées en oncologie, et sont publiés dans le New England Journal of Medicine.** Par ailleurs, l'intérêt des études en « vraie vie » est minimisé pour ces signatures et les autres alors que leur intérêt est notable et est même celui qui a permis d'asseoir l'intérêt de certains facteurs pronostiques et/ou prédictifs historiques. Les données des registres prospectifs tels que Raster pour Mammaprint® (Int J Cancer 2013;133:929-36), les données des bases SEER (Breast Cancer Res Treat 2017;163:303-31), CLALIT (Breast Cancer 2017;3 : article 32) et DCBG (J Clin Oncol 2018; 36 :735-740) n'ont pas été prises en compte alors qu'elles donnent des informations cliniques et pronostiques dans des conditions « actuelles » qui pourraient être utilement comparées avec les données des essais cliniques ou des essais prospectifs évalués de façon rétrospective (TransSATaC, ABCSG).

R1

Les objectifs et la population définis dans l'expertise ne sont pas pertinents, écartant des patientes ayant une incertitude du bénéfice à la chimiothérapie de façon inappropriée. Il n'est pas logique de vouloir écarter certaines patientes de la population à risque intermédiaire sur la base de marqueurs non validés par des études prospectives randomisées (Ki67 ou infiltrat lymphocytaire, emboles). Rappelons qu'aucun essai de désescalade de cette envergure n'a été réalisé en oncologie sur des marqueurs bio-pathologiques auparavant, et que leur intérêt limité est régulièrement rapporté dans les conférences de consensus nationales ou internationales.

À l'inverse, le groupe de travail affirme que nous n'avons pas besoin de marqueurs dans des risques intermédiaires bas, ce qui est inexact car la variabilité et la discordance des choix proposés aux femmes dans ces situations sont bien documentées. C'est précisément dans cette population, bien représentée dans les essais MINDACT et TAYLORx notamment, que la valeur prédictive des signatures (de non-intérêt de la chimiothérapie) peut être très utile en termes de désescalade. Par ailleurs, des données en situation réelle existent pour Prosigna® (Danish study), pour Mammaprint® (étude RASTER) et pour Oncotype Dx® (Clalit, SEER), portant sur plusieurs dizaines de milliers de patientes et reproduisant en termes de répartition de scores de rechute et de survie, des données très proches de celles retrouvées dans les essais cliniques disponibles ou dans les études rétrospectives. Ces cohortes collectées prospectivement devraient être incluses dans les analyses, aussi bien pour la valeur pronostique que prédictive de ces signatures.

Les conclusions du rapport nous paraissent ambiguës, puisqu'à la fois très négatives dans l'analyse des données scientifiques d'un niveau de preuve I, mais ménageant d'un autre côté la perspective de poursuivre une évaluation globale (p73).

*Celle-ci est aujourd'hui financée par la DGOS au titre des de l'innovation (RIHN) et cette pérennisation serait questionnable dans un cadre de recherche, car la DDGOS n'a pas pour objectif de financer le développement de ces outils mais d'en définir la véritable place. Le groupe de travail reconnaît lui-même (p 73) qu'arrivant au terme des grandes évaluations prospectives et randomisées (plus de 10 ans et près de 17000 patientes), il n'y a rien à attendre dans un futur proche des données scientifiques. **Il n'est donc pas logique et onéreux de proposer de maintenir ces coûts d'évaluations au titre de l'innovation via la DGOS.** Les études de recherches ou d'évaluations françaises proposées sont aujourd'hui irréalisables compte tenu des délais et moyens nécessaires, et n'amèneraient rien. Il est illusoire de penser que de nouveaux essais randomisés vont être développés en raison du recul nécessaire pour évaluer les résultats et du très grand nombre de patientes nécessaires. Les propositions du groupe de travail sur ces évaluations ou recherches en France sont déconcertantes, la France étant très en retrait de ces innovations. Les coûts publics seraient élevés pour des évaluations qui n'auraient aucun intérêt aujourd'hui.*

Nous redoutons qu'un médecin qui a obligation de moyens soit pris en défaut dans des situations où le choix de la chimiothérapie serait plus éclairé par des informations génomiques de valeur pronostique ou prédictive objectives et validées, que par d'autres éléments non validés ou subjectifs suggérés par ce rapport. Nous pensons qu'un avis aussi négatif sur les signatures qui ont aujourd'hui franchi un cap déterminant pour leur utilité clinique, expose les patientes à une perte de chance. Il est en décalage avec la communauté scientifique internationale. La parution d'un tel rapport expose les autorités de santé française à un risque considérable de recours de patientes en France, puisque certains référentiels étrangers voisins sont déjà en train de se positionner favorablement comme en Allemagne (ITWIG), au Royaume Uni (Nice) ou en Espagne (SEOM).

De façon très étonnante, les recommandations d'experts internationaux (ASCO, ESMO, Saint Gallen...), incluant des experts français, admettent un intérêt clinique pour certaines de ces signatures, mais ne font pourtant pas l'objet d'une discussion approfondie dans le rapport.

Enfin, la 8ème édition du TNM/AJCC/IUCC étant la référence pour la classification internationale des cancers en stade selon des paramètres pronostiques et prédictifs, à intégrer les signatures moléculaires des cancers du sein (AJCC 8th Edition, pgs 617, 621, 624), et là encore sans allusion particulière du groupe de travail.

Le rapport présenté, rejetant les signatures génomiques et ne souhaitant pas intégrer des données majeures de désescalade pour les malades, assume donc une exception française qui pourrait être cautionnée par une Autorité Administrative Indépendante, et qui s'affranchirait des données de la science en se positionnant au-dessus des expertises et recommandations internationales. Il sera bien difficile d'expliquer cette particularité française aux patientes que nous traitons à l'heure où nos concitoyens commencent à prendre conscience de l'intérêt de la médecine préventive et prédictive et à demander leur intégration dans les choix stratégiques de traitement.

Analyse critique du rapport d'évaluation

OBJECTIF : POPULATION CIBLE

1) Le Groupe de travail de l'HAS (GT) a louablement cherché à définir une population de patientes ayant une indication de CT basée sur des CCP de mauvais pronostics qui pourraient être candidates à une désescalade thérapeutique basée sur la géno-

mique.

Dans la population d'intérêt (T1-2, RE+, Her 2-, N0), le GT a positionné, le cut off de 1 cm, avançant l'idée, qu'en dessous de cette taille, la question d'une CT ne se posait pas pour la population d'oncologues français. Il y aurait sur ce point un consensus. En l'absence de consensus, il suffirait d'évaluer correctement le cas pour arriver à une décision de traitement.

Cette assertion est inexacte. À titre d'exemple, le dernier référentiel IGR-Curie (observé par de nombreux centres) recommande une CT ou l'emploi d'une SG pour les T1b de haut risque (définis par le grade, la présence d'embolies lymphatiques ou un Ki67 élevé).

C'est cette sous-population T1b à haut risque qui est incluse dans TAILORx .

(for 5 mm-1.0 cm plus unfavorable histological features, defined as an intermediate or poor nuclear and/or histologic grade, or lymphovascularinvasion). Rappelons qu'aucun de ces deux marqueurs n'a fait l'objet de consensus, ni d'étude de prédiction de l'intérêt de la chimiothérapie prospective randomisée, même si une analyse poolée récente positive a été récemment publiée (Denkert et al. Lancet Oncol, 2018). Il n'est donc pas logique d'exclure ces patients d'une analyse de désescalade. Le groupe de la SFMPP recommande donc fortement de réintroduire les patientes T1b dans l'analyse.

2) Le GT part du postulat que la population de TAILORx, pour les patientes RS 0-25, constitue une population à bas-risque, non représentative de la population traitée en France.

Le groupe de la SFMPP ne souscrit pas à cette affirmation. D'une part, l'absence de registres ne permet pas d'en juger. D'autre part, des éléments forts contredisent cette assertion.

Le groupe de la SFMPP s'appuie en particulier sur les données recueillies dans l'essai Européen MINDACT :

- Les critères d'inclusion dans l'essai MINDACT étaient larges et incluaient toutes les patientes non métastatiques, HER2-/HER 2+, RH+/ RH- T1-T2, N0/N1.
- La population cible de MINDACT dans lequel de nombreuses patientes Françaises ont été incluses paraît représentative de la population cible Française.
- Une large majorité des patientes incluses dans MINDACT (63 % des patientes) correspondait aux critères de la population cible de l'essai TAILORx (T1-2, N0, RH+, HER 2-).
- 69 % des patientes de ce groupe N0, RH+, Her 2- étaient considérées à bas risque clinique (tel que défini par AOL modifié) dans l'essai MINDACT, soit un pourcentage similaire à celui observé dans l'essai TAILORx (68 %).
- En outre, il faut souligner que 17,2% des tumeurs analysées dans TAILORx correspondaient à des lésions de grade 3.

MINDACT incluait des tumeurs à haut et très haut risque (triple négatives, HER2+ et N+). 28,8% des tumeurs dans MINDACT correspondaient à des lésions de grade 3.

Ces données paraissent cohérentes et on ne peut donc retenir comme valide l'assertion du GT de non représentativité des populations de TAILORx et MINDACT.

De plus, le groupe de la SFMPP tient à souligner l'importance de l'enjeu de santé publique que représente ce groupe de patientes T1-2, N0, RH+, HER 2-, dit à bas risque.

Les dernières estimations publiées de l'INCa estiment le nombre de nouveaux cas de cancer du sein en France à 59000. Les données du réseau FRANCIM (dernières

années évaluées : 2009 - 2012) montrent que 65 % des patientes < 75 ans, traitées initialement par chirurgie, seraient T1-T2 / N0. Parmi elles, environ 2/3 sont RH+. Ce qui montre que cette population représente 45% des cancers du sein traités par chirurgie première, soit près de 20 000 cancers invasifs.

Nous ne possédons pas de renseignements sur les caractéristiques histologiques et biologiques de ces tumeurs. Mais il n'y a pas lieu de penser que celles-ci sont fondamentalement différentes de celles utilisées dans l'analyse de MINDACT.

Le groupe de la SFMPP estime donc la population T1-T2 N0, RH+, Her 2-, pour laquelle se pose la question de l'utilisation d'une signature génomique dans les conditions de l'essai TAILORx, à plus de 15000 patientes en France chaque année.

3) Validité et reproductibilité des CCP

La qualité des anatomo-pathologistes et des biologistes en France est mise en exergue puis évoque la rigueur des contrôles de qualité (p41 du rapport). Le groupe de la SFMPP souscrit volontiers à la première affirmation, mais émet des réserves quant à la seconde.

Pour ce qui concerne deux des CCP retenus par le GT: Grade SBR et présence d'emboles vasculaires, il s'agit d'éléments classiques et centraux de l'analyse morphologique du cancer du sein pour lesquels il existe des référentiels de bonnes pratiques, mais pas d'évaluation externe de la qualité organisée sur le territoire français ou ailleurs.

La reproductibilité du grade SBR reste modérée, en particulier pour les grades intermédiaires et le paramètre de l'anisocaryose. Ainsi, l'essai Allemand planB, met en évidence une faible reproductibilité du Grade SBR (en particulier des Grades 2 et 3). 185/521 (35,5%) des grades 3 estimés en laboratoires locaux sont dévalués en grade 2 ou 1 par des structures de références. Ces discordances peuvent également être le reflet de l'hétérogénéité tumorale, les échantillons envoyés dans les essais cliniques n'étant pas toujours ceux analysés par le pathologiste local.

La SFMPP note que l'évaluation de la présence d'emboles vasculaires et/ou lymphatique est complexe avec un risque potentiel d'erreur. De plus, tel que noté par le GT, les emboles vasculaires dites "extensives" sont retenues comme critère majeur: il n'y a pas de littérature définissant clairement ce caractère « extensif » ni soutenant de façon formelle à ce jour cela. Le caractère extensif des emboles n'est pas défini clairement dans la littérature.

La possibilité de surévaluation du grade et de la présence d'emboles constitue donc un sujet de préoccupation. En effet, si les conclusions du pré-rapport étaient intégralement retenues (interdiction des SG pour les patientes qui ont 2 critères de gravité ou une tumeur de grade 3- p 42 du rapport-), elles amèneraient inéluctablement un sur-traitement pour certaines patientes.

D'une façon générale, il est surprenant qu'un GT devant travailler sur l'expertise de signatures génomiques fasse des propositions de traitement sur d'autres éléments (sans niveau de preuve 1). Cela paraît hors sujet, puisqu'il ne s'agit pas de la mission principale du GT de faire des recommandations de prise en charge.

STRATEGIE DE RÉFÉRENCE UTILISANT LES CCP POUR EVALUER L'APPORT

D'UNE SG

Préambule

Le GT indique qu'aucune étude n'a été retrouvée démontrant l'utilité d'une association SG-CCP. Cette affirmation n'est pas vraie, et l'expression pourrait paraître tendancieuse.

À titre d'exemple, la seule publication évoquée (mais dont la référence bibliographique n'est pas fournie dans le document) cherchait à établir l'intérêt d'une introduction de données cliniques (score RSPC) dans un modèle de prédiction établi à partir de la SG seule. Le score RSCP (Oncotype-données cliniques) dégrade le modèle de prédiction pour l'appréciation du bénéfice d'une chimiothérapie adjuvante. Il pourrait cependant avoir un intérêt pour les scores intermédiaires

Ce qui a été montré pour ONCOTYPE n'est donc pas l'absence d'intérêt d'une SG en appoint de CCP, mais plutôt l'absence d'intérêt de l'appoint de CCP dans un modèle de prédiction, établi à partir de la SG seule. Pour les signatures EndoPredict, (EPICLIN) et Prosigna, l'intégration de données cliniques semble améliorer leurs performances.

Stratégie de référence définie par le GT

1) D'après le GT, les référentiels antérieurs à 2009 définissaient le risque et les indications en fonction unique des CCP, sans mention d'usage aux SG. Ces référentiels sont donc les seuls susceptibles d'être utilisés afin de définir la stratégie de référence à même d'évaluer de l'impact de l'introduction des SG.

Ainsi, le GT promeut une stratégie « pure » dite de référence, basée sur l'évaluation des risques, établie à partir de 5 CCP majeurs de mauvais pronostic : l'âge < 40 ans, la taille tumorale ($\geq T2$), le grade 3, la présence d'embolies et la présence de macro-métastases ganglionnaires.

Dans la stratégie de référence promue par le GT, la présence de deux critères induit systématiquement la prescription d'une CT adjuvante. Dans certains centres, le seul grade 3 suffit à cette prescription.

Le GT ne retient pas 4 CCP utilisés dans de nombreux référentiels : l'appréciation de la prolifération par le Ki67, le compte mitotique, le niveau d'expression des récepteurs à la progestérone et le niveau d'expression des récepteurs à l'Estradiol. En revanche, le GT les juge à même d'aider à la décision, sans en préciser la manière.

2) Le groupe de la SFMPP note que le GT, pour définir la stratégie de référence, s'appuie sur les conclusions des conférences de consensus de St Gallen 2007 (a) et 2009 (b) et de Nice St Paul en 2011 (c).

Le groupe de la SFMPP a donc exploré de manière attentive ces 3 références :

a) La conférence de St Gallen de 2007 définissait les CCP et les seuils d'intervention de la manière suivante :

- Âge ≤ 35 ans ou > 35 ans.*
- Niveau d'expression des récepteurs à l'Estradiol ET à la progestérone (et pas seulement l'état d'expression +ou - des récepteurs à l'Estradiol). La confé-*

rence de St Gallen ne fixait pas pour l'un ou l'autre groupe de récepteurs de valeurs-seuils.

- Grade SBR 1, 2 ou 3.
- Prolifération évaluée par : le Ki67. 3 états : Bas <14%, intermédiaire 14-30, élevé > 30 et/ou l'index mitotique : niveaux non précisés.
- Emboles vasculaires : présence ou absence.
- Taille tumorale : > 5 cm, 2.1-5 cm < 2cm.
- Atteinte ganglionnaire : N0, N1, N2

b) La conférence de St Gallen 2009 reprenait les mêmes CCP, mais en ajoutant cependant les SG (if available), comme élément principal pour la décision dans les cas difficiles.

Pour la conférence de St Gallen 2009, le résultat des SG était considéré comme contributif pour les scores bas (indication d'hormonothérapie seule) et pour les scores élevés (indication d'une association CT-Hormonothérapie).

c) À côté de cette mauvaise lecture de la conférence de St Gallen 2009, le groupe de la SFMPP a constaté un second manquement grave dans l'analyse : il n'existe aucune trace de recommandations émises par la conférence de consensus de Nice St-Paul 2011 à propos des traitements systémiques adjuvants.

En lieu et place, le groupe de la SFMPP a trouvé une simple veille bibliographique. Le groupe de la SFMPP s'étonne de cette seconde confusion qui nuit grandement à la prise en compte de son propos.

3) Afin d'aider à la compréhension des propositions du GT, le groupe de la SFMPP propose donc d'utiliser les éléments de la conférence de St Gallen 2007.

Comme on peut aisément le constater, la conférence de consensus citée propose d'utiliser des CCP et des seuils d'intervention différents de ceux promus par le GT.

Ainsi, il apparaît que la prise en compte des 5 CCP retenus par le GT, la manière de les utiliser, de les additionner et de définir les seuils d'intervention est postérieure à l'introduction des SG dans la conférence de consensus de St Gallen.

On notera également que cette stratégie établie par le GT ne trouve sa place dans aucune autre conférence de consensus.

Enfin, sans procéder à une revue exhaustive des référentiels français, on constate qu'elle ne correspond ni aux propositions du référentiel IGR-CURIE (qui introduit largement les SG, utilise le Ki67 et/ou le compte mitotique dans son algorithme) ni au référentiel de l'APHP (qui ne retient pas les emboles vasculaires dans son algorithme, mais les cite simplement comme élément d'intérêt). Pour illustrer son propos, le groupe de la SFMPP en reprend les éléments principaux et complète par l'appréciation du risque selon Adjuvant Online modifié.

Appréciation du « risque clinique » par l'algorithme Adjuvant Online modifié

- pour les tumeurs grade 1 : taille tumorale > 3 cm,
- pour les tumeurs grade 2 : taille tumorale > 2 cm,
- pour les tumeurs grade 3 : taille tumorale > 1 cm.

Référentiel APHP : appréciation du haut risque

- Tumeurs grade 1 :
 - Taille tumorale ≥ 3 cm
 - Lorsque la tumeur est < 3cm et que le Ki67 est élevé, la décision est à

prendre au cas par cas par la RCP.

- Tumeurs grade 2 :
 - Taille tumorale ≥ 2 cm et Ki67 élevé (ou prolifération forte).
 - Taille ≥ 3 cm dans les autres cas.
- Tumeurs grade 3 :
 - Taille tumorale ≥ 1 cm et Ki67 élevé (ou prolifération forte)

Lorsque le Ki67 est bas, la décision est à prendre au cas par cas par la RCP.

Le référentiel IGR-CURIE

Quatre niveaux de risque sont définis :

- Toutes les tumeurs T1a (≤ 5 mm) sont à très bas risque.
- Les tumeurs T1b (5-10 mm) sont à bas risque si de Grade 1 et prolifération faible, à risque intermédiaire dans les autres cas.
- Les prises en charge prévues sont les suivantes :
 - très bas risque : abstention thérapeutique
 - bas risque : hormonothérapie exclusive,
 - haut risque : chimiothérapie et hormonothérapie,
 - le risque intermédiaire pourrait amener la prescription d'un test génomique afin d'aider à la décision. Le type de test à utiliser n'est pas précisé.
- Le risque intermédiaire y est défini des manières suivantes :
 - Tumeurs de grade 1 :
 - Tumeurs T2 (T1 : bas risque).
 - Tumeurs de grade 2 :
 - Tumeurs T1c (1-2 cm) + Ki67 > 20 et / ou présence d'embolies et/ou lymphatiques (ILV); (T1c + Ki67 < 20 ET sans embolies : bas risque).
 - T2 quel que soit le Ki67, qu'il y ait ou non des embolies.
 - Tumeurs de grade 3 :
 - Tumeurs T1 (T2 : haut risque).

Le Ki67 et la présence d'embolies ne sont pas des critères utilisés pour la sélection des tumeurs de grade 3.

Plusieurs constats s'imposent :

1) Dans les référentiels APHP et IGR-CURIE l'adjonction du Ki67 afin d'affiner le modèle n'est, de l'avis du GT, pas pertinente en raison de la non reproductibilité de l'examen, de la haute fréquence des résultats intermédiaires et des seuils décisionnels non consensuels.

2) Si le référentiel IGR-CURIE positionne nettement les embolies vasculaires dans son modèle de décision, ce n'est pas le cas pour le référentiel de l'APHP. La présence d'embolies tumorales figure comme élément à prendre en compte, mais n'est pas reprise dans l'algorithme de décision. À propos de l'utilisation de la présence d'embolies vasculaires comme CCP majeur, le Groupe de la SFMPP remarque que son usage est très inégalement répandu et a fluctué dans le temps. À titre d'exemple, le référentiel IGR-CURIE 2014-2015 ne les prenait pas en compte dans son algorithme.

3) Le GT indique que le risque clinique tel qu'il est évalué par l'algorithme Adjuvant Online modifié n'est plus utilisé par les cliniciens. Cependant, si, comme le préconise le GT, l'évaluation du Ki67 n'est pas retenue comme CCP, on constate que le référentiel APHP s'approche très fortement de l'appréciation du risque clinique.

Pour des CCP identiques, les 2 référentiels les plus utilisés d'Île de France amènent

souvent à poser des indications différentes de traitement adjuvant. De même, les référentiels français et internationaux n'ont jamais pu dégager de consensus et ont utilisé de manière fluctuante dans le temps, des CCP variables avec des seuils d'intervention flous.

Évaluation de la stratégie de référence élaborée par le GT

La SFMPP salue les efforts du GT afin d'établir une stratégie de référence. Notons que cette stratégie n'est actuellement pas retenue par de nombreux centres (IGR-CURIE, APHP). Elle n'est ni validée (aucun essai prospectif ni même évaluation rétrospective), ni consensuelle (en France ou à l'International), ni répandue. Basée sur des CCP peu reproductibles, non utilisables pour la prise de décision d'un traitement adjuvant, la stratégie de référence du GT est, de plus, récente et reflète presque uniquement une prise de position du GT.

Le groupe de la SFMPP rappelle cependant que le GT n'a pas été mandaté pour une proposition de ce type et que les diverses sociétés professionnelles et savantes n'ont pas été sollicitées à ce titre.

La stratégie de référence proposée par le GT ne peut donc, en aucune manière, servir de base à l'évaluation de stratégies plus anciennes, validées par deux essais de phase III. Elle ne saurait pas davantage être utilisée à l'élaboration d'une réponse à la sollicitation de l'HAS et encore moins à un encadrement de la prescription des SG en fonction des CCP.

Méthodes d'évaluation de l'utilité clinique des SG

Le groupe de la SFMPP estime que la stratégie de référence du GT est inappropriée pour évaluer l'utilité clinique de SG.

- En revanche, le groupe constate que toutes les expériences de mise à disposition des SG aux praticiens se sont soldées par des bouleversements de pratique. Si l'on étudie, par exemple les données de l'observatoire PONDx, un taux de changement des décisions de RCP est observé dans plus de 40% des cas avec une diminution globale du recours à la chimiothérapie de 35% des cas. Les données issues de ces observatoires ne sauraient être négligées. Rajoutées les données de Prosigna et de EndoPredict sinon le rapport sera biaisé OptiGen : 17% de changement de décision thérapeutique et 25% de discordance entre profil IHC et BM / **St Gallen. Adendom 35,8% de changement de thérapie 7,5% ajout de la CT et 28,4% retrait.**

Validité des données issues de TAILORx

TAILORx est un essai de phase III, de grande ampleur, indépendant de l'industrie, bien mené, dont les conclusions sont probantes (Niveau de Preuves Ia).

Cet essai vient compléter des données concordantes recueillies au cours de MINDACT.

TAILORx s'inscrit dans une très longue liste de publications portant sur des essais, des observatoires et des registres. Toutes ces publications indiquent une même direction : les SG sont à même d'isoler, de manière originale et indépendante des CCP, une population pour laquelle une chimiothérapie adjuvante n'apporte pas de bénéfice. TAILORx a permis d'en préciser davantage les seuils d'intervention.

TAILORx et préconisations du GT

Les données de TAILORx permettent d'évaluer les conséquences des préconisations du GT (avec ses restrictions d'indication sur les SG). Ces préconisations aboutiraient potentiellement à :

- Sous-traiter 188/1259 (15%) des T1b à risque.
- Sur-traiter au moins 58/582 (10%) des patientes ≤ 40 ans (Recurrence Score < 11)
- Sur-traiter 2445/9716 (25%) des patientes T2
- Sur-traiter les patientes de grade SBR3 avec un RS < 11 (111/1676) et une grande part des 884/1676 (52%) patientes avec un RS compris entre 11 et 26.
- Un nombre non évaluable, mais consistant, de tumeurs présentant des embolies. Le registre de la National Data Base permet de mettre en relation Recurrence score (RS) et présence d'embolies. Pour les années 2010-2012 : 1682/9199 (9%) des tumeurs avec embolies avaient un RS < 11 ; 875/12062 (7,2%) des tumeurs avec embolies avaient un RS entre 11 et 25x ; 1118/6704 (16,6%) des tumeurs avec embolies avaient un RS > 25 .

Ces nombres, rapportés à la population française, sont considérables.

Dans les données prospectives, il faut également citer, outre CLALIT et SEER, les données de la DANISH study (JCO 2017 Prosigna). Dans l'étude du Danish Breast Cancer Cooperative Group database de 2000 à 2003 : 2,558 RH+/HER2- dont 1,395 N+ qui confirme la valeur pronostique de Prosigna® du ROR et des types moléculaires. Il n'y a pas de données prospectives pronostiques de grande échelle disponibles pour EndoPredict mais les études sont en cours.

Réponses de l'association de patients « Vivre comme avant »

LISIBILITÉ ET CLARTÉ

L1

Auriez-vous des commentaires à effectuer pour améliorer la lisibilité et la clarté du document provisoire qui vous a été adressé ?

Réponse :
non

POINT DE VUE DE VOTRE ORGANISME

P1

Auriez-vous des remarques à formuler quant à votre appropriation du travail d'analyse effectué par la HAS ?

Réponse :
Pas de remarque

P2

Votre organisme peut-il éclairer la HAS sur les éventuelles conséquences pratiques et/ou organisationnelles de la présente évaluation pour vos membres/adhérents ?

Réponse :
non

REMARQUES LIBRES

R1

Avez-vous d'autres commentaires à formuler ?

Réponse :
Pas de commentaire

Rapport d'évaluation de l'utilité clinique des signatures génomiques dans le cancer du sein de stade précoce

Annexe 13. Recherche documentaire

La recherche a porté sur les publications en langue anglaise et français.

Les équations de recherche ci-dessous présente de façon synthétique les étapes successives de cette interrogation dans la base de données Medline.

Le nombre total de références obtenues par interrogation des bases de données bibliographiques est 259, veille bibliographique incluse.

Interrogation de base de données

Type d'étude / sujet

Termes utilisés

| Signatures d'expression multigénique dans le cancer précoce du sein | |
|--|--|
| Essais contrôlés randomisés, études comparatives, études prospectives | 01/2002 – 05/2018 |
| Étape 1 | (Oncotype Dx OR Mammaprint OR Endopredict OR Prosigna OR 21 gene recurrence score* OR 21 gene expression* OR EP score* OR 8 gene expression* OR EPclin OR 70-gene signature* OR 70 gene expression* OR PAM50)/ti,ab OR (Oncotype Dx OR Mammaprint OR Endopredict OR Prosigna OR 21 gene recurrence score* OR 21 gene expression* OR EP score* OR EPclin OR 70 gene signature* OR 70 gene expression* OR PAM50)/ot |
| ET | |
| Étape 2 | (metaanalys* OR meta-analys* OR meta analysis OR systematic review* OR systematic overview* OR systematic literature review* OR systematical review* OR systematical overview* OR systematical literature review* OR systematic literature search OR pooled analysis OR random* OR clinical trial* OR comparative stud* OR versus OR cohort* OR longitudinal stud* OR follow-up stud* OR prospective stud*)/ti OR (Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies OR Cohort Studies OR Longitudinal Studies OR Follow-Up Studies OR Prospective Studies)/de OR (meta-analysis OR randomized controlled trial OR Clinical Trial OR comparative study)/pt OR cochrane database syst rev/ta |

de : descriptor ; de maj : major topic ; ti : title ; ab : abstract ; ta : journal title ; pt : publication type ; ot : other term ; ! : explosion du terme générique

Sites internet consultés

Dernière consultation : mars 2018

Adelaide Health Technology Assessment – AHTA
Australia and New Zealand Horizon Scanning Network
California Technology Assessment Forum – CTAF
Canadian Agency for Drugs and Technologies in Health – CADTH
Centre fédéral d'expertise des soins de santé – KCE
Centre for Clinical Effectiveness – CCE
Centre for Reviews and Dissemination databases
Comité d'Évaluation et de Diffusion des Innovations Technologiques – CEDIT
Health Services Technology Assessment Text – HSTAT
Horizon Scanning Research & Intelligence Centre
Institute for Clinical Evaluative Sciences – ICES
Institut national d'excellence en santé et en services sociaux – INESSS
Institute for Health Economics Alberta – IHE
Institute for Quality and Efficiency in Health Care – IQWiG
International Network of Agencies for Health Technology Assessment – INAHTA
Malaysian Health Technology Assessment Section – MaHTAS
McGill University Health Centre

Medical Services Advisory Committee – MSAC
National Coordinating Centre for Health Technology Assessment – NCCHTA
National Institute for Health and Clinical Excellence – NICE
NHS Evidence
Ontario Health Technology Advisory Committee – OHTAC
Queensland Government Health Policy Advisory Committee on Technology
Swedish Agency for Health Technology Assessment and Assessment of Social Services – SBU
Tripdatabase
West Midlands Health Technology Assessment Collaboration – WMHTA

Veille bibliographique :

En complément, une veille a été réalisée jusqu'à fin juillet 2018 sur Medline à partir des équations décrites ci-dessus.



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