

AVIS

Pertinence de l'utilisation des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée

Une production de l'Institut national
d'excellence en santé
et en services sociaux (INESSS)

Direction de l'évaluation et de la pertinence
des modes d'intervention en santé



Pertinence de l'utilisation des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée

Rédaction

Mélanie Béland

Léon Nshimyumukiza

Collaboration

Frédéric Breton

Éric Potvin

Coordination scientifique

Jim Boulanger

Direction

Michèle de Guise

Le rapport a été présenté au Comité d'excellence clinique en santé de l'Institut national d'excellence en santé et en services sociaux (INESSS) lors de sa réunion du 11 septembre 2020.

Le contenu de cette publication a été rédigé et édité par l'INESSS.

Membres de l'équipe de projet

Auteurs

Mélanie Béland, Ph. D.
Léon Nshimyumukiza, M. Sc., Ph. D.

Collaborateurs internes

Frédéric Breton, B. Sc.
Éric Potvin, Ph. D.

Coordonnateur scientifique

Jim Boulanger, Ph. D.

Adjoint à la direction

Yannick Auclair, Ph. D.

Direction des services de santé

Michèle de Guise, M.D., FRCPC

Repérage d'information scientifique

Renaud Lussier, M.S.I.
Mathieu Plamondon, M.S.I.
Lysane St-Amour, M.B.S.I.

Gestion de l'information

Mike Benigeri, Ph. D.
El-Hadji Malick Ndao, M. Sc.

Soutien administratif

Lolita Haddad

Équipe de l'édition

Denis Santerre
Hélène St-Hilaire
Nathalie Vanier

Sous la coordination de

Renée Latulippe, M.A.

Avec la collaboration de

Littera Plus, révision linguistique
Mark A. Wickens, traduction

Dépôt légal

Bibliothèque et Archives nationales du Québec, 2021
Bibliothèque et Archives Canada, 2021
ISBN 978-2-550-88223-7 (PDF)

© Gouvernement du Québec, 2021

La reproduction totale ou partielle de ce document est autorisée à condition que la source soit mentionnée.

Pour citer ce document : Institut national d'excellence en santé et en services sociaux (INESSS). Pertinence de l'utilisation des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée. Avis rédigé par Mélanie Béland et Léon Nshimyumukiza. Québec, Qc : INESSS; 2021. 104 p.

L'Institut remercie les membres de son personnel qui ont contribué à l'élaboration du présent document.

Comité consultatif

Pour ce rapport, les membres du comité d'experts sont :

D^r Régen Drouin, généticien, CHUL et Centre mère-enfant Soleil (CHU de Québec–Université Laval)

D^r Zaki El Haffaf, généticien, CHUM

D^r Pierre-Hugues Fortier, ORL/Chirurgie cervicofaciale, CHUS (CIUSSS de l'Estrie)

D^r Jean-François Ouellet, chirurgien oncologue, Hôtel-Dieu de Québec (CHU de Québec – Université Laval)

D^r Richard Payne, ORL/Chirurgie cervicofaciale, CUSM

D^r Marc Pusztaszeri, pathologiste, Hôpital général juif (CIUSSS du Centre-Ouest-de-l'Île-de-Montréal)

D^r Keith Richardson, ORL/Chirurgie cervicofaciale, CUSM

Comité d'évolution des pratiques en oncologie (CEPO)

Président

D^r Félix Couture, président, hématologue et oncologue médical, Hôtel-Dieu de Québec (CHU de Québec – Université Laval)

Vice-président

D^r Ghislain Cournoyer, vice-président, hématologue et oncologue médical, Hôpital régional de Saint-Jérôme (CISSS des Laurentides)

Membres

M^{me} Karine Almanric, pharmacienne, Hôpital de la Cité-de-la-Santé (CISSS de Laval)

D^r Jean-Sébastien Aucoin, hématologue et oncologue médical, Centre hospitalier affilié universitaire régional (CIUSSS de la Mauricie-et-du-Centre-du-Québec)

M. Philippe Bouchard, pharmacien, Hôpital Maisonneuve-Rosemont (CIUSSS de l'Est-de-l'Île-de-Montréal)

D^r Alexis Bujold, radio-oncologue, Hôpital Maisonneuve-Rosemont (CIUSSS de l'Est-de-l'Île-de-Montréal)

D^r Normand Gervais, chirurgien oncologue, Centre hospitalier régional du Grand-Portage (CISSS du Bas-Saint-Laurent)

M^{me} Marie-Pascale Guay, pharmacienne, Hôpital général juif (CIUSSS du Centre Ouest-de-l'Île-de-Montréal)

D^r Bernard Lespérance, hématologue et oncologue médical, Hôpital du Sacré-Cœur de Montréal (CIUSSS du Nord-de-l'Île-de-Montréal)

M^{me} Nathalie Letarte, pharmacienne, Centre hospitalier de l'Université de Montréal (CHUM), représentante du Programme de gestion thérapeutique des médicaments

D^r Ari Meguerditchian, chirurgien oncologue, Site Glen (CUSM)

D^r Jean-François Ouellet, chirurgien oncologue, Hôtel-Dieu de Québec (CHU de Québec – Université Laval)

D^r Benoît Samson, hématologue et oncologue médical, Hôpital Charles-Le Moyne (CISSS de la Montérégie-Centre)

D^r François Vincent, radio-oncologue, Centre hospitalier affilié universitaire régional (CIUSSS de la Mauricie-et-du-Centre-du-Québec)

Comité d'excellence clinique en services de santé

Président

M. Daniel La Roche, directeur de l'évaluation, de la qualité, de l'éthique, de la planification et des affaires juridiques, CHU de Québec – Université Laval

Vice-président

M. Serge Dumont, vice-président, professeur émérite, Faculté des sciences sociales, Université Laval

Membres

M^{me} Danielle Boucher, infirmière praticienne spécialisée en néphrologie, CHU de Québec – Université Laval

D^r Luigi Lepanto, radiologue; directeur des services professionnels, Direction des affaires médicales et universitaires, Centre hospitalier de l'Université de Montréal; professeur titulaire de clinique, Faculté de médecine, École de santé publique, Université de Montréal

M^{me} Aude Motulsky, pharmacienne et professeure adjointe, École de santé publique de l'Université de Montréal (ESPUM); chercheuse, Centre de recherche du Centre hospitalier de l'Université de Montréal (CR CHUM)

M. Thomas Poder, professeur adjoint, Département de gestion, évaluation et politique de santé, École de santé publique de l'Université de Montréal, Université de Montréal; chercheur, Centre de recherche de l'Institut universitaire en santé mentale de Montréal; professeur associé, Département d'économie et Département de médecine de famille et de médecine d'urgence, Université de Sherbrooke

D^r Martin Potter, médecin de famille, CISSS de la Montérégie-Ouest, GMF de Vaudreuil-Soulanges, professeur adjoint clinique au programme de résidence de médecine familiale de l'Université de Montréal

D^{re} Charo Rodriguez, professeure titulaire, Département de médecine familiale, Université McGill

D^{re} Nadia Roumeliotis, médecine de soins intensifs de l'enfant, CHU Sainte-Justine; professeure adjointe de clinique, Université de Montréal

D^{re} Elise Sirois-Giguère, chirurgienne générale; professeure adjointe d'enseignement clinique, Faculté de médecine et des sciences de la santé, Université de Sherbrooke

M^{me} Stéphanie Therrien, directrice de l'éthique et du développement durable, Revenu Québec

M. Pierre-Yves Therriault, ergothérapeute; professeur et directeur, Département d'ergothérapie, Université du Québec à Trois-Rivières

M^{me} Nathalie Thiffault, conseillère-cadre en soins infirmiers aux continuums de soins critiques, cardiologie et neurologie, CIUSSS de la Mauricie-et-du-Centre-du-Québec

D^{re} Chantal Vallée, interniste, CISSS de la Montérégie-Centre; professeure agrégée, Faculté de médecine et des sciences de la santé, Université de Sherbrooke

Membres citoyens

M. Pierre-Luc Désilets

M. Guy Poulin

Comité scientifique permanent des analyses de biologie médicale (CSABM)

Président

D^r François Rousseau, médecin biochimiste, CHU de Québec – Université Laval

Vice-président

D^r Lambert Busque, hématologue, CIUSSS de l'Est-de-l'Île-de-Montréal

Membres

D^r Guy Fink, biochimiste clinique, CIUSSS de l'Estrie, CHUS

D^r Cédric Yansouni, microbiologiste infectiologue, Centre universitaire de santé McGill

Membre citoyenne

M^{me} Suzanne K. Bédard

Déclaration d'intérêts

Les auteurs de l'avis déclarent n'avoir aucun conflit d'intérêts ou de rôle relatif à l'objet de l'évaluation; aucun financement externe n'a été obtenu pour sa réalisation.

Aucun conflit d'intérêts n'a été relevé chez les experts consultés et les membres des divers comités consultés.

Responsabilité

L'Institut assume l'entière responsabilité de la forme et du contenu définitifs du présent document. Les conclusions et recommandations ne reflètent pas forcément les opinions des lecteurs externes ou celles des autres personnes consultées pour les besoins du présent dossier.

TABLE DES MATIÈRES

RÉSUMÉ.....	I
SUMMARY.....	VII
SIGLES ET ACRONYMES	XII
GLOSSAIRE	XIV
INTRODUCTION.....	1
1. MÉTHODOLOGIE	9
1.1. Questions d'évaluation	9
1.2. Cadre conceptuel.....	9
1.3. Stratégie de repérage de l'information scientifique	10
1.4. Critères et processus de sélection des études.....	11
1.5. Extraction des données	12
1.6. Évaluation de la qualité méthodologique des études	12
1.7. Méthode de contextualisation	12
1.8. Méthodes d'analyse et de synthèse des données	13
1.9. Méthodes d'analyse économique	14
1.10. Processus de validation scientifique.....	14
1.11. Processus d'appréciation de la preuve scientifique et formulation des recommandations	15
1.12. Gestion des conflits d'intérêts.....	15
2. RÉSULTATS	17
2.1. Description des études repérées.....	17
2.2. Performance des tests moléculaires commercialisés pour l'identification des nodules bénins à cytologie indéterminée	17
2.2.1. Proportion de résultats moléculairement bénins (BCR).....	18
2.2.2. Sensibilité et spécificité	18
2.2.3. Valeurs prédictives	18
2.2.4. Prévalence de la maladie parmi les nodules à cytologie indéterminée	19
2.3. Performance analytique des tests moléculaires disponibles	21
2.4. Évaluation du nombre de chirurgies pratiquées et du risque résiduel de malignité à la suite d'un résultat moléculairement bénin	22
2.4.1. Afirma GEC	23
2.4.2. Impact de l'introduction du test Afirma GEC	27
2.4.3. Afirma GSC	30
2.4.4. ThyroSeq v2.1	32
2.4.5. ThyroSeq v3	34
2.5. Comparaison des nodules à cytologie indéterminée moléculairement bénins avec les nodules à cytologie bénigne	35
2.6. Portrait de la prise en charge actuelle des nodules thyroïdiens à cytologie indéterminée	38
2.6.1. Recommandations internationales.....	38

2.6.2. Contexte québécois.....	42
2.7. Perspective des patients.....	45
2.8. Consultation - Perspective des experts	48
2.9. Volet économique	48
2.9.1. Littérature scientifique	48
2.9.2. Évaluation économique primaire.....	52
2.9.3. Analyse d'impact budgétaire	57
DISCUSSION.....	62
RÉSUMÉ DES DÉLIBÉRATIONS ET RECOMMANDATIONS	69
RÉFÉRENCES.....	73
ANNEXE A.....	80
Stratégie de repérage de l'information scientifique	80
ANNEXE B.....	81
Algorithme décisionnel	81
ANNEXE C.....	85
Sélection des études	85
ANNEXE D.....	90
Résultats de l'évaluation de la qualité méthodologique des études	90
ANNEXE E.....	96
Impact de la prévalence de la maladie.....	96
Validité analytique	97
ANNEXE F.....	98
Estimation de la population admissible, analyses de scénarios pour l'impact budgétaire et estimation des coûts unitaires	98
ANNEXE G.....	103
Gradation de la preuve	103

LISTE DES TABLEAUX

Tableau 1	Indications d'une biopsie par aspiration à l'aiguille fine selon les caractéristiques échographiques.....	2
Tableau 2	Système de classification Bethesda.....	3
Tableau 3	Interprétation du résultat de l'analyse ThyroSeq selon la version employée.....	5
Tableau 4	Critères d'inclusion et d'exclusion des études scientifiques.....	11
Tableau 5	Performance des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée	20
Tableau 6	Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat Afirma GEC bénin	25
Tableau 7	Impact de l'analyse Afirma GEC sur le taux de chirurgies thyroïdiennes et la prévalence de cas malins chez les patients présentant un nodule à cytologie indéterminée	29

Tableau 8	Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat Afirma GSC bénin	31
Tableau 9	Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat ThyroSeq v2.1 bénin	33
Tableau 10	Prise en charge des nodules à cytologie indéterminée GEC bénins en comparaison avec celle des nodules à cytologie bénigne.....	36
Tableau 11	Recommandations en lien avec la prise en charge des nodules thyroïdiens à cytologie indéterminée.....	40
Tableau 12	Préférence des patients pour la prise en charge des nodules à cytologie indéterminée selon le risque de malignité, le coût et les propriétés diagnostiques des tests moléculaires	45
Tableau 13	Impact de l'évaluation moléculaire sur la qualité de vie des patients	47
Tableau 14	Évaluation moléculaire et rapport coût-efficacité en fonction du risque échographique	49
Tableau 15	Rapport coût-efficacité des différentes approches diagnostiques	51
Tableau 16	Intrants cliniques et économiques considérés dans le modèle	54
Tableau 17	Résultats du scénario de base de l'analyse coût-utilité de ThyroSeq v3 versus la lobectomie	55
Tableau 18	Résultats du scénario de base de l'analyse coût-utilité d'Afirma GSC versus la lobectomie	55
Tableau 19	Répartition des patients selon le scénario.....	58
Tableau 20	Impact budgétaire associé au remboursement de ThyroSeq v3	59
Tableau 21	Impact budgétaire associé au remboursement d'Afirma GSC	60

LISTE DES FIGURES

Figure 1	Prise en charge privilégiée des nodules thyroïdiens selon le résultat de la BAAF	10
Figure 2	Volume de biopsies par aspiration à l'aiguille fine pratiquées par établissement en 2018 au Québec*	42
Figure 3	Nombre d'actes enregistrés au Québec entre 2009 et 2018 selon le type de chirurgie thyroïdienne (sans égard au résultat de la cytologie)	44
Figure 4	Modèle d'analyse de l'efficience.....	53
Figure 5	Nuage de points ThyroSeq v3 vs lobectomie.....	56
Figure 6	Nuage de points Afirma GSC vs lobectomie.....	57

RÉSUMÉ

Introduction

Le nodule thyroïdien est une pathologie relativement fréquente, mais maligne dans une minorité de cas. La biopsie par aspiration à l'aiguille fine (BAAF) est généralement utile pour distinguer les nodules bénins de ceux qui sont malins et orienter la prise en charge vers l'observation ou la résection chirurgicale. Dans 20 à 30 % des cas cependant, le résultat de la cytologie est indéterminé (Bethesda III ou IV) et la prise en charge optimale est plus difficile à déterminer. Dans ces circonstances, le seul moyen d'établir un diagnostic définitif consiste à réséquer le nodule. Les données disponibles indiquent qu'un peu plus de la moitié des nodules à cytologie indéterminée font l'objet d'une résection. Toutefois, l'examen histologique des spécimens réséqués a révélé la présence d'un nodule bénin dans environ 75 % des cas. Par conséquent, cette pratique est parfois perçue comme inutile, en plus de présenter un risque non négligeable d'hypothyroïdie postopératoire qui nécessite une hormonothérapie de substitution.

Les tests moléculaires ont été développés dans le but d'exclure la malignité sans recourir à la chirurgie. Ces tests (ThyroSeq v3, Afirma GSC et ThyGeNext/ThyraMIR) sont offerts à un coût élevé par des laboratoires privés établis aux États-Unis. Selon certaines sociétés savantes, les nodules à cytologie indéterminée pourraient éviter la chirurgie et demeurer en observation si l'analyse moléculaire pouvait prédire un diagnostic bénin avec un risque résiduel de maladie inférieur ou égal à 5 % (valeur prédictive négative [VPN] ≥ 95 %).

La Direction de la biovigilance et de la biologie médicale (DBBM) de la Direction générale des affaires universitaires, médicales, infirmières et pharmaceutiques (DGAUMIP) du ministère de la Santé et des Services sociaux (MSSS) a demandé à l'INESSS d'évaluer la pertinence de recourir aux tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée.

Le présent avis a pour objectif de faciliter la décision du demandeur : 1) en évaluant la performance des analyses moléculaires disponibles pour prédire le diagnostic histopathologique des nodules bénins (validité clinique et analytique); 2) en évaluant la capacité de ces tests à réduire le nombre de chirurgies inutiles; 3) en examinant l'aspect sécuritaire associé à la surveillance active des nodules à cytologie indéterminée qui sont moléculairement bénins; 4) en dressant un portrait de la prise en charge actuelle des nodules thyroïdiens à cytologie indéterminée; et 5) en estimant l'impact économique associé au remboursement de ce type d'analyse.

Méthodes

Une revue rapide de la littérature scientifique a été réalisée afin de documenter les divers aspects en lien avec la performance et l'utilité clinique des tests moléculaires, la perspective des patients et les données d'efficacité publiées par d'autres pays ou provinces. Le portrait québécois de la prise en charge actuelle des nodules thyroïdiens a été dressé au moyen des données expérientielles et contextuelles issues de la

consultation d'experts québécois et de la base de données SMOD de la Régie de l'assurance maladie du Québec (RAMQ). Une analyse économique a également été réalisée au moyen de ces données.

Résultats

Deux des trois analyses moléculaires présentement disponibles sur le marché américain ont fait l'objet d'une publication des données de validation clinique. Le taux de résultats moléculairement bénins (BCR), qui correspond à la population susceptible d'éviter la chirurgie, a été de 61 % avec ThyroSeq v3 et de 54 % avec Afirma GSC. Des valeurs similaires de sensibilité (> 90 %) et de valeur prédictive négative (> 95 %) ont été observées pour les deux analyses alors que la spécificité a été de 82 % avec ThyroSeq v3 et de 68 % avec Afirma GSC. Étant donné que la VPN varie en fonction de la prévalence de la maladie dans la population dans laquelle le test est réalisé, le risque résiduel en présence d'un résultat négatif varie lui aussi selon le contexte. Le théorème de Bayes prédit que la chirurgie peut être évitée de façon sécuritaire (avec un risque résiduel ≤ 5 %) si la prévalence locale de la maladie est inférieure à 40 % lorsque le résultat est généré avec l'analyse ThyroSeq v3, et inférieure à 29 % lorsque la version Afirma GSC est employée. Mais, dans un contexte où l'utilisation des tests moléculaires contribue à limiter le nombre de nodules réséqués, la valeur de la prévalence locale de la maladie est désormais inconnue. Différentes approches ont été employées par les auteurs dans le but d'estimer le risque résiduel de maladie à la suite d'un résultat moléculairement bénin. Ces approches ont toutefois généré des résultats imprécis, soit parce que les nodules non réséqués ont été exclus de l'équation ou parce que tous les nodules non réséqués ont été considérés d'emblée comme histologiquement bénins.

La revue de la littérature a indiqué que le recours à la catégorie « indéterminée » pour caractériser le résultat de la cytologie est très variable d'une étude à l'autre. Une certaine variation dans l'interprétation du résultat de la cytologie par différents cytopathologistes a également été documentée. Ainsi, l'expérience de chaque établissement de santé en matière d'examen cytopathologique sur la biopsie par aspiration à l'aiguille fine de la thyroïde pourrait influencer sur le contenu de la catégorie indéterminée (prévalence de la maladie) et, par conséquent, sur le risque résiduel des nodules moléculairement bénins non réséqués. Au Québec, l'analyse des données contextuelles a relevé la présence d'écarts importants quant au volume annuel de BAAF compilées pour chaque établissement de santé. Par conséquent, le risque résiduel de maladie à la suite d'un résultat moléculairement bénin est susceptible de varier grandement d'un établissement à l'autre.

Bien que la prévalence de la maladie ait une incidence sur la performance locale des tests pour l'identification des nodules bénins, les experts consultés se sont dits peu préoccupés par cette variable et son effet sur le risque résiduel de maladie lorsque le résultat est moléculairement bénin. Selon eux, le suivi clinique/échographique réservé aux nodules moléculairement bénins, couplé à la progression lente des lésions malignes de stade précoce, permettraient d'anticiper que les nodules faussement négatifs qui auraient échappé à la détection moléculaire seraient décelés, le cas échéant, sans que

cela porte un important préjudice aux patients concernés. Néanmoins, le suivi des nodules moléculairement bénins devrait être fait en exerçant une étroite surveillance pour assurer l'identification précoce des nodules faussement négatifs qui pourraient évoluer vers des cancers agressifs. Le cas des patients perdus de vue au suivi clinique représente une préoccupation dans ce contexte.

Le taux de chirurgies thyroïdiennes parmi les nodules à cytologie indéterminée soumis à l'évaluation moléculaire a été recensé dans le but de préciser l'utilité clinique de chaque test (selon la version évaluée). En l'absence de comparaison adéquate avec la pratique actuelle (sans test), les données compilées pour l'ensemble des nodules évalués avec chaque test ont révélé un taux de résection de 51 %, 27 %, 44 % et 44 % avec les tests Afirma GEC, Afirma GSC et ThyroSeq v2.1 et v3, respectivement. Selon les données des mêmes études, la proportion de patients susceptibles d'éviter la chirurgie sur la base du ratio BCR (*Benign call rate*) a été de 44 % avec Afirma GEC, 68 % avec Afirma GSC, 74 % avec ThyroSeq v2.1 et 58 % avec ThyroSeq v3. Le taux de nodules moléculairement bénins réséqués a été de 14 %, 6 %, 33 % et 7 % avec Afirma GEC, Afirma GSC, ThyroSeq v2.1 et v3, respectivement.

Globalement, le niveau de preuve associé à la démonstration de l'utilité clinique est faible. Les données publiées pour les versions présentement disponibles sont limitées. Parmi les études repérées, le groupe expérimental, composé de nodules soumis à l'évaluation moléculaire, n'a pas fait l'objet d'une comparaison directe avec un groupe représentant la pratique standard et constitué de nodules aux propriétés similaires évalués au cours de la même période. Dans plusieurs études, les critères guidant l'usage du test étaient inconnus ou absents (usage systématique); lorsque l'usage était sélectif, les critères appliqués pour la sélection étaient rarement énoncés. Par ailleurs, comme la durée médiane du suivi était rarement précisée par les auteurs, il a été généralement difficile de savoir si celle-ci avait été suffisamment longue pour permettre de repérer tous les nodules réséqués à la suite de l'évaluation moléculaire. Ainsi, un doute persiste quant aux retombées réelles de ces analyses dans la population ciblée. Les experts étaient cependant d'avis que le besoin de santé lié au diagnostic des nodules thyroïdiens à cytologie indéterminée est important et que le remboursement des analyses moléculaires évaluées aurait un impact clinique et organisationnel favorable.

Données économiques

Sur la base des hypothèses formulées, lesquelles sont assorties d'incertitude, la probabilité de l'efficacité de l'approche moléculaire pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée serait variable, en comparaison avec la lobectomie. Au seuil de 50 000 \$/QALY (*quality-adjusted life year*) gagnée, l'emploi de ThyroSeq v3 serait probablement efficace (69 %), alors que l'efficacité associée à l'utilisation du test Afirma GSC demeure incertaine (52 %).

L'impact budgétaire associé au remboursement public des tests moléculaires serait aussi variable selon les scénarios envisagés. D'après le scénario de base, le remboursement public de l'analyse ThyroSeq v3 pourrait diminuer les coûts de 0,960 M\$ (0,173 - 5,6 M\$) sur 3 ans comparativement à la lobectomie diagnostique. Cependant, le remboursement du test pourrait augmenter les coûts lorsque :

- le risque de malignité pour un nodule à cytologie indéterminée est supérieur à 32,5 %;
- la proportion de nodules moléculairement suspects réséqués est supérieure à 87 %;
- la proportion de nodules moléculairement bénins réséqués dépasse 18,8 %;
- le coût unitaire du test est supérieur à 5 715 \$.

Le remboursement public d'Afirma GSC pourrait, quant à lui, augmenter les coûts de 0,677 M\$ (0,083 M\$ - 5,5 M\$) sur 3 ans comparativement à la lobectomie diagnostique. Cependant, le remboursement du test pourrait diminuer les coûts lorsque :

- le risque de malignité pour un nodule à cytologie indéterminée est inférieur à 18,2 %;
- la spécificité de l'analyse Afirma GSC est supérieure à 73,3 %;
- la proportion de nodules moléculairement suspects réséqués est inférieure à 71,5 %;
- la proportion de nodules moléculairement bénins réséqués est inférieure à 8,2 %;
- le coût unitaire du test est inférieur à 5 152 \$.

Résumé des délibérations

Malgré l'incertitude relative à la validité et à l'utilité clinique de ces tests, les membres du comité délibératif (CEC santé et CSABM) en ont reconnu le caractère prometteur. Les membres sont d'avis que la couverture publique de ces tests pourrait constituer une décision juste et raisonnable, mais uniquement en présence de modalités qui tiendraient compte du contexte d'incertitude.

Motifs de la position :

- l'emploi de ce type de test, en remplacement de la lobectomie diagnostique, permettrait au patient d'éviter une chirurgie comportant un risque d'hypothyroïdie postopératoire et libérerait du temps opératoire dans ce secteur d'activité ainsi que pour d'autres spécialités;
- un résultat moléculairement bénin pourrait constituer un facteur rassurant pour le patient, en comparaison avec le risque initial de maladie associé aux nodules à cytologie indéterminée;

- la progression lente des lésions malignes de stade précoce et le suivi réservé aux nodules moléculairement bénins permettraient d'envisager une prise en charge sécuritaire, même en cas de résultat faussement négatif;
- une économie de coûts est envisageable pour le diagnostic et la prise en charge des nodules à cytologie indéterminée si l'efficacité anticipée de l'approche moléculaire est avérée;
- certains membres ont exprimé un malaise concernant le remboursement public de ce type d'analyse en raison de la difficulté rencontrée pour reconnaître la valeur de l'approche moléculaire en l'absence d'une étude de validation clinique non financée par le fabricant;
- des études de validation clinique indépendantes du fabricant ne sont pas attendues à court terme;
- il n'y a pas d'étude confirmant l'utilité clinique de ces tests en comparaison avec la pratique standard.

Recommandations

L'Institut national d'excellence en santé et en services sociaux (INESSS) reconnaît le caractère prometteur des tests ThyroSeq v3 et Afirma GSC pour préciser le diagnostic des nodules thyroïdiens dans certains cas de cytologie indéterminée. Toutefois, compte tenu de l'incertitude liée à la validité et à l'utilité clinique de ces tests et de l'effet de certaines variables sur les coûts de l'approche moléculaire, l'INESSS est d'avis que leur couverture publique pourrait constituer une décision juste et raisonnable si les conditions suivantes sont respectées :

- L'utilisation devrait être restreinte aux patients chez qui le résultat du test est susceptible d'influer sur la prise en charge.

Les nodules admissibles devraient avoir les caractéristiques suivantes :

- nodule à cytologie de catégorie Bethesda III (après une 2^e BAAF) ou Bethesda IV; et
- dont la taille est comprise entre 1 et 4 cm; et
- qui présentent des facteurs de risque cliniques ou échographiques intermédiaires (équivoques ou discordants).

Les patients pour lesquels une chirurgie est indiquée (facteurs de risque élevé), ceux pour qui la lobectomie diagnostique n'est pas envisagée en raison d'un faible risque de malignité, de même que ceux qui manifestent une préférence pour la chirurgie ou l'observation, ne devraient pas se voir offrir une évaluation moléculaire.

- La mise en œuvre préalable d'une étape de validation clinique en contexte québécois¹.
- Une entente est conclue avec le fabricant pour partager le risque financier.

Autres éléments :

- Le patient doit être informé du risque résiduel de malignité associé à la décision de ne pas réséquer un nodule à cytologie indéterminée moléculairement bénin avant d'opter pour une analyse moléculaire.
- Les patients dont le nodule est moléculairement bénin devraient faire l'objet d'un suivi tous les 12 à 24 mois durant 4 à 5 ans afin d'assurer l'identification précoce des nodules faussement négatifs qui pourraient évoluer vers des cancers agressifs.
- L'INESSS suggère la réalisation d'une collecte de données prospectives afin de compiler l'information relative 1) aux caractéristiques de base des nodules sélectionnés pour l'évaluation moléculaire (Bethesda III ou IV, taille, caractéristiques échographiques), 2) au choix du test (ThyroSeq v3 ou Afirma GSC) et au résultat de l'analyse (bénin ou suspect) et 3) à la décision de prise en charge subséquente (observation ou chirurgie). Les données recueillies pourront ensuite faire l'objet d'une évaluation au terme de laquelle la révision des critères d'admissibilité ou un guide d'usage optimal pourrait être proposé.

¹ La confirmation indépendante de la validité clinique implique 1) que tous les nodules à cytologie indéterminée soumis au fabricant (selon les critères d'admissibilité mentionnés plus haut) soient aussi réséqués (dans le contexte de validation) pour que la performance de l'approche moléculaire soit établie sur la base du diagnostic histopathologique et 2) que les résultats issus de cette démonstration soient minimalement comparables à ceux générés dans le cadre de l'étude de validation clinique originale.

SUMMARY

Advisability of using molecular tests to clarify the diagnosis of thyroid nodules with indeterminate cytology

Introduction

Thyroid nodules are relatively common but malignant in only a minority of cases. Fine-needle aspiration biopsy (FNAB) is generally useful for distinguishing benign from malignant nodules and guiding management to either observation or surgical resection. However, in 20 % to 30 % of cases, the cytology result is indeterminate (Bethesda III or IV) and the optimal management more difficult to determine. In such circumstances, the only way to make a definitive diagnosis is to resect the nodule. The available data indicate that just over half of cytologically indeterminate nodules are resected. However, histological examination of resected specimens revealed a benign nodule in approximately 75 % of the cases. Therefore, this practice is sometimes seen as unnecessary, in addition to entailing a significant risk of postoperative hypothyroidism requiring hormone replacement therapy.

Molecular tests have been developed to rule out malignancy without resorting to surgery. These tests (ThyroSeq v3, Afirma GSC and ThyGeNext/ThyraMIR) are offered at a high cost by private laboratories in the United States. According to certain expert groups, surgery for cytologically indeterminate nodules could be avoided and they could remain under observation if the molecular test could predict a benign diagnosis with a residual risk of disease of 5 % or less (negative predictive value [NPV] ≥ 95 %).

The Direction de la biovigilance et de la biologie médicale (DBBM) of the Direction générale des affaires universitaires, médicales, infirmières et pharmaceutiques (DGAUMIP) of the Ministère de la Santé et des Services sociaux (MSSS) asked INESSS to assess the advisability of using molecular tests to clarify the diagnosis of thyroid nodules with indeterminate cytology.

The purpose of this report is to facilitate the requestor's decision: 1) by evaluating the performance of the available molecular tests for predicting the histopathologic diagnosis of benign nodules (clinical and analytical validity); 2) by assessing the ability of these tests to reduce the number of unnecessary surgeries; 3) by examining the safety of observation for cytologically indeterminate nodules with a benign molecular test result; 4) by providing a picture of the current management of thyroid nodules with indeterminate cytology; and 5) by estimating the economic impact of covering this type of test.

Methods

A rapid review of the scientific literature was conducted to document the various aspects related to the performance and clinical utility of the molecular tests, the patient perspective, and the efficiency data published in other countries and provinces. The

picture of current thyroid nodule management in Québec was constructed from experiential and contextual data from consultations with Québec experts and the Régie de l'assurance maladie du Québec (RAMQ)'s SMOD database. These data were also used to perform an economic analysis.

Results

Clinical validation data have been published for two of the three molecular tests currently available on the U.S. market. The benign call rate (BCR), which corresponds to the population likely to avoid surgery, was 61 % with ThyroSeq v3 and 54 % with Afirma GSC. Similar figures for sensitivity ($> 90\%$) and negative predictive value ($> 95\%$) were observed for both tests, while specificity was 82 % for ThyroSeq v3 and 68 % for Afirma GSC. Since the NPV varies according to the prevalence of the disease in the population in which the test is performed, the residual risk when the result is negative also varies according to the context. Bayes' theorem predicts that surgery can be safely avoided (with a residual risk $\leq 5\%$) if the local disease prevalence is less than 40 % when the result is obtained with ThyroSeq v3 and less than 29 % when Afirma GSC is used. However, in a context where the use of molecular tests helps limit the number of nodules resected, the local prevalence rate for the disease is unknown. Many authors have used different approaches to estimate the residual risk of disease following a benign molecular test result. These approaches, however, yielded imprecise results, either because unresected nodules were excluded from the equation or because all unresected nodules were considered histologically benign from the outset.

The literature review indicated that the use of the "indeterminate" category to characterize the cytology result varies widely from study to study. Some variation in the interpretation of the cytology result by different cytopathologists was also documented. Thus, each healthcare facility's experience in the cytopathologic examination of thyroid fine-needle aspirates could influence the contents of the indeterminate category (disease prevalence) and, consequently, the residual risk of unresected molecularly benign nodules. In Québec, the analysis of the contextual data revealed significant differences in the annual number of FNABs compiled for each healthcare facility. Therefore, the residual risk of disease following a benign molecular test result is likely to vary greatly from one facility to another.

Although disease prevalence has an impact on the tests' local performance for identifying benign nodules, the experts consulted expressed little concern over this variable and its effect on the residual risk of disease in the case of a benign molecular test result. In their view, with the clinical/ultrasonographic monitoring reserved for molecularly benign nodules and with the slow progression of early-stage malignant lesions, it could be expected that any false-negative nodules that may have escaped molecular detection would be detected, if applicable, without this causing significant harm to the patients concerned. Nevertheless, molecularly benign nodules should be observed closely to ensure early identification of false-negative nodules that could progress to aggressive cancers. The case of patients lost to clinical follow-up is a concern in this regard.

The thyroid surgery rate for cytologically indeterminate nodules subjected to molecular testing was identified in order to assess the clinical utility of each test (based on the version evaluated). In the absence of an adequate comparison with current practice (no testing), the data compiled for all the nodules analyzed with each test showed resection rates of 51 %, 27 %, 44 % and 44 % for Afirma GEC, Afirma GSC and ThyroSeq v2.1 and v3, respectively. According to the data from the same studies, the proportion of patients likely to avoid surgery based on the BCR (benign call rate) was 44 % with Afirma GEC, 68 % with the Afirma GSC, 74 % with the ThyroSeq v2.1 and 58 % with ThyroSeq v3. The proportion of molecularly benign nodules resected was 14 %, 6 %, 33 % and 7 % with Afirma GEC, Afirma GSC, and ThyroSeq v2.1 and v3, respectively.

Overall, the level of evidence associated with the clinical utility demonstration is low. The published data for the currently available versions is limited. In the studies identified, the experimental group, which consisted of nodules subjected to molecular testing, was not directly compared with a group representing the standard practice and consisting of nodules with similar characteristics evaluated over the same period. In several studies, the criteria guiding the use of the test were not known or absent (systematic use). When the tests were used selectively, the selection criteria were rarely stated. In addition, because the median follow-up was rarely indicated by the authors, it was generally unclear whether the duration of follow-up was long enough to identify all the nodules resected after molecular testing. Thus, there is still some doubt as to the actual impact of these tests in the target population. However, the experts were of the opinion that the health need stemming from the diagnosis of thyroid nodules with indeterminate cytology is significant and that covering the molecular tests evaluated would have a favourable clinical and organizational impact.

Economic data

Based on the assumptions made, which are accompanied by some uncertainty, the probability of the molecular approach being efficient in clarifying the diagnosis of thyroid nodules with indeterminate cytology is variable compared to lobectomy. At the threshold of \$50,000/QALY (quality-adjusted life year) gained, the use of ThyroSeq v3 would probably be efficient (69 %), while the efficiency associated with the use of Afirma GSC is uncertain (52 %).

The budget impact associated with public coverage of the molecular tests would also vary according to the scenario considered. In the base-case scenario, public coverage of ThyroSeq v3 could reduce costs by \$0,960 M (\$0,173 - \$5,6 M) over 3 years relative to diagnostic lobectomy. However, coverage of this test could increase costs when:

- The risk of malignancy for a cytologically indeterminate nodule is greater than 32,5 %;
- The proportion of molecularly suspicious nodules resected is greater than 87 %;
- The proportion of molecularly benign nodules resected is greater than 18,8 %;
- The unit cost of the test is greater than \$5 715.

Public coverage of Afirma GSC could increase costs by \$0,677M (\$0,083M - \$5,5M) over 3 years relative to diagnostic lobectomy. However, coverage of this test could decrease costs when:

- The risk of malignancy for a cytologically indeterminate nodule is less than 18,2 %;
- The specificity of Afirma GSC is greater than 73,3 %;
- The proportion of molecularly suspicious nodules resected is less than 71,5 %;
- The proportion of molecularly benign nodules resected is less than 8,2 %;
- The unit cost of the test is less than \$5 152.

Summary of deliberations

Despite the uncertainty regarding the validity and clinical utility of these tests, the members of the deliberative committee (CEC – santé and CSABM) recognize the promise they hold. The members feel that public coverage of these tests could be a fair and reasonable decision, but only if there are terms and conditions that take this uncertainty into account.

Reasons for the position:

- The use of this type of test as an alternative to diagnostic lobectomy would enable the patient to avoid surgery entailing a risk of postoperative hypothyroidism and would free up surgical time in this sector of activity and for other specialties;
- A molecularly benign result could be a reassuring factor for the patient compared to the initial risk of disease associated with cytologically indeterminate nodules;
- With the slow progression of early-stage malignant lesions and the monitoring reserved for molecularly benign nodules, one could anticipate safe management, even in the event of a false-negative result;
- Cost savings in terms of diagnosing and managing nodules with indeterminate cytology can be expected if the anticipated efficiency of the molecular approach is established;
- Some members expressed uneasiness about public coverage of this type of test because of the difficulty recognizing the value of the molecular approach in the absence of a non-manufacturer-funded clinical validation study;
- Manufacturer-independent clinical validation studies are not expected in the short term;
- There are no studies confirming the clinical utility of these tests compared to standard practice.

Recommendations

The Institut national d'excellence en santé et en services sociaux (INESSS) recognizes the promise held by ThyroSeq v3 and Afirma GSC for clarifying the diagnosis of thyroid nodules in certain cases of indeterminate cytology. However, given the uncertainty regarding these tests' validity and clinical utility and the impact of certain variables on the costs of the molecular approach, INESSS feels that public coverage of these tests could constitute a fair and reasonable decision if the following conditions are met:

- Their use should be limited to patients whose test result is likely to influence management.

Eligible nodules should have the following characteristics:

- Bethesda III (after a 2nd FNAB) or Bethesda IV category cytology; and
- 1 to 4 cm in size; and
- Exhibit intermediate clinical or ultrasonographic risk factors (equivocal or discordant).

Molecular testing should not be proposed to patients for whom surgery is indicated (high-risk factors), those for whom diagnostic lobectomy is not being considered because of a low risk of malignancy, or those who express a preference for surgery or observation.

- The prior implementation of a clinical validation step in the Québec context².
- An agreement with the manufacturer to share the financial risk is concluded.

Other items:

- The patient must be informed of the residual risk of malignancy associated with the decision not to resect a cytologically indeterminate nodule with a benign molecular test result before opting for molecular testing.
- Patients with a molecularly benign nodule should be followed every 12 to 24 months for 4 to 5 years to ensure early identification of any false-negative nodules that could progress to aggressive cancers.
- INESSS suggests that prospective data be gathered to compile information regarding 1) the basic characteristics of nodules selected for molecular testing (Bethesda III or IV, size, ultrasonographic characteristics), 2) the choice of test (ThyroSeq v3 or Afirma GSC) and the test result (benign or suspicious) and 3) the subsequent management decision (observation or surgery). These data can then be examined, after which a review of the eligibility criteria or an optimal use guide could be proposed.

² Independent confirmation of clinical validity implies 1) that all cytologically indeterminate nodules submitted to the manufacturer (as per the above-mentioned eligibility criteria) are also resected (in the validation context) so that the performance of the molecular approach is established on the basis of the histopathologic diagnosis and 2) that the results of this demonstration are at least comparable to those obtained in the initial clinical validation study.

SIGLES ET ACRONYMES

AACE	American Association of Clinical Endocrinologists
ACE	American College of Endocrinology
ADN	Acide désoxyribonucléique
AHNS	American Head and Neck Society
AME	Associazione Medici Endocrinologi
ARNm	Acide ribonucléique messenger
ATA	American Thyroid Association
AUS/FLUS	<i>Atypia of undetermined significance/follicular lesion of undetermined significance</i> ; atypie de signification indéterminée/lésion folliculaire de signification indéterminée
BAAF	Biopsie par aspiration à l'aiguille fine
BCR	<i>Benign call rate</i> ; proportion de résultats moléculairement bénins
BTA	British Thyroid Association
CEC	Comité d'excellence clinique
CEPO	Comité de l'évolution des pratiques en oncologie
Chx	Chirurgie
CNA	<i>Copy number alteration</i> ; altération du nombre de copies
CSABM	Comité scientifique permanent des analyses de biologie médicale
DBBM	Direction de la biovigilance et de la biologie médicale
DGAUMIP	Direction générale des affaires universitaires, médicales, infirmières et pharmaceutiques
Dx	Diagnostic
ETA	European Thyroid Association
FN	Faux négatif
FN-ISG	Faux négatif- <i>incomplete surgical group</i> ; incluant les nodules bénins non réséqués
FN/SFN	<i>Follicular neoplasm/suspicious for a follicular neoplasm</i> ; néoplasme folliculaire/suspicion d'un néoplasme folliculaire
GEA	<i>Gene expression alteration</i> ; altération de l'expression génique
GEC	<i>Gene expression classifier</i>
GF	<i>Gene fusion</i> ; fusion génique
GSC	<i>Genomic sequencing classifier</i>
INESSS	Institut national d'excellence en santé et en services sociaux
MSSS	Ministère de la Santé et des Services sociaux

NCCN	National Comprehensive Cancer Network
NIFTP	<i>Noninvasive follicular thyroid neoplasm with papillary-like nuclear features</i> ; néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire
NLR	Nerf laryngé récurrent
QALY	<i>Quality-adjusted life year</i>
RAMQ	Régie de l'assurance maladie du Québec
SNV/I	<i>Single-nucleotide variant/indel</i> ; variations nucléotidiques simples/ insertions ou délétions
TI-RADS	<i>Thyroid Imaging, Reporting and Data System</i>
TSH	Thyréostimuline
UPMC	Centre médical de l'Université de Pittsburgh
VPN	Valeur prédictive négative
VPP	Valeur prédictive positive

GLOSSAIRE

Analyse coût-utilité

Analyse économique qui permet de comparer des interventions avec des effets de nature différente en tenant compte de la qualité de vie et en employant une seule mesure, soit l'utilité [Crochard-Lacour et LeLorier, 2000].

Arbre décisionnel

Représentation graphique des options et résultats possibles utilisée en analyse décisionnelle [INESSS, 2018].

Benign call rate (BCR)

Proportion de nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test moléculaire [Ohuri *et al.*, 2019].

Faux négatif (FN)

Nodule au résultat moléculairement bénin qui s'avère malin à l'examen histopathologique du spécimen réséqué.

Taux de faux négatifs : rapport du nombre de résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules moléculairement bénins réséqués.

Taux de faux négatifs incluant les nodules bénins non réséqués (FN-ISG) : rapport du nombre de résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules qui ont obtenu un résultat bénin au test moléculaire, bien que le nombre de nodules bénins réséqués soit incomplet au moment de l'analyse [Harrell *et al.*, 2019].

Hémithyroïdectomie (ou lobectomie)

Intervention chirurgicale qui consiste à extraire la moitié de la glande thyroïde (lobe droit ou gauche) ainsi que sa portion centrale, l'isthme³.

Modèle de Markov

Modélisation quantitative comprenant un ensemble particulier d'états de santé mutuellement exclusifs et complets, pour lesquels il existe des probabilités de transition d'un état à un autre, y compris la probabilité de rester dans le même état. Les états ont généralement une durée uniforme et les probabilités de transition ne changent pas avec le temps [INESSS, 2018].

³ Société canadienne du cancer (SCC). Chirurgie du cancer de la thyroïde [site Web]. Disponible à : <https://www.cancer.ca/fr-ca/cancer-information/cancer-type/thyroid/treatment/surgery/?region=on>.

Prévalence de la maladie

Proportion des nodules à cytologie indéterminée dont le diagnostic histopathologique final est malin ou de type « néoplasme thyroïdien folliculaire non invasif avec caractéristique nucléaire de type papillaire » (NIFTP) [Cibas et Ali, 2017; Haugen *et al.*, 2017].

Sensibilité

Caractéristique de la performance d'un test diagnostique, qui se définit comme la proportion des personnes qui ont un résultat de test positif parmi les malades; elle se calcule ainsi : $[\text{vrais positifs} \div (\text{vrais positifs} + \text{faux négatifs})]$ [INESSS, 2018].

Spécificité

Caractéristique de la performance d'un test diagnostique, qui se définit comme la proportion des personnes qui ont un résultat de test négatif parmi les non-malades; elle se calcule ainsi : $[\text{vrais négatifs} \div (\text{vrais négatifs} + \text{faux positifs})]$ [INESSS, 2018].

Valeur prédictive négative (VPN)

Caractéristique de la performance d'un test, qui se définit comme la proportion des personnes qui n'ont pas la maladie parmi celles qui ont un résultat négatif à un test diagnostique; elle se calcule ainsi : $[\text{vrais négatifs} \div (\text{vrais négatifs} + \text{faux négatifs})]$.

Remarque : La VPN varie selon la prévalence de la maladie dans la population étudiée [INESSS, 2018].

Valeur prédictive positive (VPP)

Caractéristique de la performance d'un test, qui se définit comme la proportion des personnes qui ont la maladie parmi celles qui ont un résultat positif à un test diagnostique; elle se calcule ainsi : $[\text{vrais positifs} \div (\text{vrais positifs} + \text{faux positifs})]$. La VPP varie en fonction de la prévalence de la maladie dans la population étudiée [INESSS, 2018].

INTRODUCTION

La détection d'un nodule thyroïdien est chose courante [Durante *et al.*, 2018; Haugen *et al.*, 2016]. Dans les régions du monde qui bénéficient d'un apport suffisant en iode, environ 5 % de la population présente un nodule palpable à l'évaluation clinique. L'examen de la région du cou par imagerie diagnostique pour des motifs non liés à la thyroïde a également permis l'identification fortuite d'un nodule asymptomatique, souvent petit, chez 19 à 68 % des individus. La prévalence des nodules thyroïdiens et le taux de multinodularité augmentent avec l'âge, le genre féminin et l'indice de masse corporelle. Néanmoins, les nodules thyroïdiens sont généralement bénins. Le suivi clinique vise à exclure la présence d'une tumeur maligne, observée dans 7 à 15 % des cas, pour laquelle un traitement chirurgical est habituellement recommandé [Haugen *et al.*, 2016].

Le cancer de la thyroïde

Les principaux types de cancer de la thyroïde incluent les carcinomes différenciés (90 % papillaires, 5 % folliculaires, 2 % à cellule de Hürthle), médullaires (2 %) et anaplasiques (1 %) [Noone *et al.*, 2018]. La Société canadienne du cancer estime à 1 890 le nombre de nouveaux cas de cancer de la thyroïde diagnostiqués au Québec en 2019 (8 200 au Canada) et à environ 45 le nombre de patients qui décéderont de cette maladie (230 au Canada) [Société canadienne du cancer (SCC), 2019]. Entre 1984 et 2015, le cancer de la thyroïde a connu la plus forte hausse annuelle de son taux d'incidence normalisé selon l'âge (4,9 % chez l'homme et 5,1 % chez la femme). Malgré cette hausse significative, le taux de mortalité est demeuré stable. Le taux prédit de survie nette à 5 et 10 ans est de 98 et 97 %, respectivement. L'introduction de nouvelles techniques d'imagerie diagnostique couplée à une surveillance médicale accrue seraient liées à la hausse observée du taux d'incidence du cancer de la thyroïde [Topstad et Dickinson, 2017; Vaccarella *et al.*, 2016].

La biopsie par aspiration à l'aiguille fine

Les anomalies palpables ainsi que celles repérées incidemment lors d'examens radiologiques font habituellement l'objet d'une évaluation échographique suivie ou non (selon le risque de malignité) d'une biopsie par aspiration à l'aiguille fine (BAAF; tableau 1). D'après l'American Thyroid Association (ATA), la BAAF est recommandée pour les nodules ≥ 1 cm qui présentent des caractéristiques échographiques à risque élevé ou intermédiaire de malignité ainsi que pour les nodules $\geq 1,5$ cm qui présentent un faible risque [Haugen *et al.*, 2016]. La BAAF peut aussi être considérée dans les cas de nodules ≥ 2 cm dont les caractéristiques échographiques révèlent un très faible risque de malignité. Depuis quelques années, plusieurs experts du Québec emploient le système TI-RADS (*Thyroid Imaging, Reporting and Data System*) pour stratifier le risque de malignité (TR1-TR5) en fonction des caractéristiques échographiques observées [Tessler *et al.*, 2017]. Les seuils de taille déterminés pour recommander la BAAF à chaque niveau de risque sont légèrement différents de ceux proposés par l'ATA (tableau 1).

Tableau 1 Indications d'une biopsie par aspiration à l'aiguille fine selon les caractéristiques échographiques

ATA [Haugen <i>et al.</i> , 2016]				
Suspicion	Caractéristiques échographiques	Risque de malignité	Indication pour la BAAF	
Élevée	Nodule solide hypoéchogène ou nodule partiellement cystique avec une composante solide hypoéchogène accompagné d'au moins une des caractéristiques suivantes : marges irrégulières (infiltrant, microlobulé), microcalcifications, dimension plus haute que large, calcifications en bordure avec une composante extrusive de tissus mous, extension extrathyroïdienne.	> 70-90 %	≥ 1 cm	
Intermédiaire	Nodule solide hypoéchogène avec des marges lisses sans microcalcification, extension extrathyroïdienne ou dimension plus haute que large.	10-20 %	≥ 1 cm	
Faible	Nodule solide isoéchogène ou hyperéchogène ou nodule partiellement cystique avec zones solides excentrées sans microcalcification, marge irrégulière ou extension extrathyroïdienne, ou dimension plus haute que large.	5-10 %	≥ 1,5 cm	
Très faible	Nodule spongiforme ou partiellement cystique en l'absence des caractéristiques échographiques décrites pour les catégories de suspicion précédentes.	< 3 %	≥ 2 cm*	
Bénigne	Nodules exclusivement cystiques (sans composante solide)	< 1 %	Non indiqué†	
ACR TI-RADS [Tessler <i>et al.</i> , 2017]				
Catégorie	Caractéristiques échographiques	Pointage cumulé (5 catégories)	Suspicion	Indication pour la BAAF
Composition	Cystique ou presque entièrement cystique : 0 point; spongiforme : 0 point; mixte cystique/solide : 1 point; solide : 2 points	≥ 7	TR5 Élevée	≥ 1 cm
Échogénicité	Anéchogène : 0 point; hyper ou isoéchogène : 1 point; hypoéchogène : 2 points; très hypoéchogène : 3 points	4-6	TR4 Modérée	≥ 1,5 cm
Catégorie	Caractéristiques échographiques	Pointage cumulé (5 catégories)	Suspicion	Indication pour la BAAF
Dimension	Plus large que haute : 0 point; plus haute que large : 3 points	3	TR3 Légère	≥ 2,5 cm
Marges	Lisses : 0 point; mal définies : 0 point; lobulées ou irrégulières : 2 points; extension extrathyroïdienne : 3 points	2	TR2 Aucune	Non indiqué
Foyers échogènes	Aucun ou artéfacts en queue de comète : 0 point; microcalcifications : 1 point; calcifications en bordure : 2 points; foyers ponctués : 3 points	0	TR1 Bénin	Non indiqué

BAAF : biopsie par aspiration à l'aiguille fine.

* L'observation sans BAAF est aussi une option raisonnable.

† L'aspiration du kyste peut être envisagée pour un drainage symptomatique ou cosmétique.

La BAAF est pratiquée par palpation ou sous guidage échographique par un spécialiste de la radiologie diagnostique, de l'otorhinolaryngologie et de la chirurgie cervicofaciale ou de l'endocrinologie. Le système de classification Bethesda (tableau 2) est communément employé pour stratifier les résultats de la cytologie de la glande thyroïde selon six catégories diagnostiques [Cibas et Ali, 2017]. Il permet d'estimer le risque de malignité et de guider les décisions de prise en charge subséquentes. Le système Bethesda est particulièrement utile pour les spécimens des catégories II, V et VI dont les valeurs prédictives négative et positive sont suffisamment élevées (VPN catégorie II : 96,3 %; VPP catégories V et VI : 75,2 % et 98,6 %) pour appuyer la surveillance active (II) ou la résection chirurgicale (V et VI) [Bongiovanni *et al.*, 2012]. Toutefois, de 20 à 30 % des résultats de cytologie sont de nature indéterminée (catégories III et IV). Dans ces cas, le risque de malignité peut varier de 10 à 30 % pour les spécimens de catégorie III et de 25 à 40 % pour ceux de catégorie IV.

Tableau 2 Système de classification Bethesda

Catégorie	Diagnostic	Risque de malignité*
I	<i>Non-diagnostic</i>	5-10 %
II	Bénin	0-3 %
III	Atypie de signification indéterminée ou lésion folliculaire de signification indéterminée	10-30 %
IV	Néoplasme folliculaire ou suspicion d'un néoplasme folliculaire	25-40 %
V	Suspicion de malignité	50-75 %
VI	Malin	97-99 %

Tiré de Cibas et Ali, 2017.

BAAF : biopsie par aspiration à l'aiguille fine.

* Incluant les néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire, ou NIFTP, dont la résection chirurgicale est nécessaire pour établir un diagnostic définitif [Haugen *et al.*, 2017].

Les nodules thyroïdiens à cytologie indéterminée

La décision de prise en charge des nodules thyroïdiens à cytologie indéterminée est évaluée dans le contexte des facteurs de risque cliniques, échographiques et cytologiques de chaque patient et selon ses préférences. Les options incluent la surveillance active (suivi clinique/échographique) et la lobectomie, une intervention chirurgicale pratiquée dans le but de préciser le diagnostic des BAAF des catégories III et IV [Haugen *et al.*, 2016].

Une méta-analyse portant sur 25 445 nodules dont 5 012 étaient à cytologie indéterminée a révélé que 39 % des résultats de cytologie de catégorie III et 70 % de ceux de catégorie IV ont fait l'objet d'une résection chirurgicale (55 % pour les deux catégories) [Bongiovanni *et al.*, 2012]. Dans la majorité des cas cependant, l'examen histopathologique a confirmé la présence d'un nodule bénin (III : 84 % et IV : 74 %). Par conséquent, cette pratique peut être perçue comme une source de surtraitement,

susceptible d'avoir des répercussions importantes sur le patient ainsi que sur les ressources du système de santé.

Les complications observées après une chirurgie thyroïdienne sont peu fréquentes. Elles incluent l'hypocalcémie (< 5 %), l'hématome (< 2 %) et la paralysie du nerf laryngé récurrent (NLR) (< 2 %), qui coordonne le mouvement des cordes vocales [Nouraei *et al.*, 2017; Weiss *et al.*, 2015]; le risque de développer ces complications est significativement moins élevé à la suite d'une lobectomie (hémithyroïdectomie) en comparaison avec la thyroïdectomie totale. Par ailleurs, une importante séquelle pouvant survenir à la suite de l'hémithyroïdectomie est l'hypothyroïdie [Kandil *et al.*, 2013; Verloop *et al.*, 2012]. Bien que la moitié de la thyroïde ait été préservée par ce type d'intervention, permettant la production endogène de thyroxine (T₄; hormone thyroïdienne), environ 20 % des patients présentent une hypothyroïdie postopératoire (permanente ou transitoire) qui nécessite une hormonothérapie de substitution. Le risque d'hypothyroïdie postopératoire peut être réduit par l'identification préalable des individus qui ont un niveau préopératoire élevé de TSH (thyroïdostimuline).

Au cours de la dernière décennie, diverses analyses moléculaires ont été développées et proposées sur le marché américain afin de préciser la nature des spécimens à cytologie indéterminée, d'améliorer la stratification du risque de malignité et de réduire le nombre de chirurgies inutiles. Ces analyses combinent différentes approches incluant la détection de mutations, d'insertions/délétions, de fusions géniques, d'altérations du nombre de copies ou de l'expression de plusieurs gènes cibles. Les algorithmes employés pour rapporter le résultat sous une forme dichotomique (bénin/négatif ou suspect/positif) demeurent la propriété exclusive des laboratoires qui ont développé et offrent ces services. Par conséquent, ces tests ne peuvent être réalisés dans les laboratoires publics québécois. De plus, comme ces tests n'ont pas été développés dans le but de les rendre disponibles sous la forme de trousse commercialisées, ils sont exemptés du processus d'évaluation de la Food and Drug Administration, tel que déclaré par chaque fournisseur dans la clause de non-responsabilité du rapport de laboratoire qui fait état du résultat de l'analyse ([ThyroSeq](#), [Afirma](#), [ThyGeNext/ThyraMIR](#)).

ThyroSeq

L'analyse ThyroSeq a été développée au Département de pathologie du Centre médical de l'Université de Pittsburgh (UPMC). La version d'origine (v0) a été conçue pour permettre l'identification de mutations ponctuelles (*BRAF*, *NRAS*, *HRAS* et *KRAS*) et de réarrangements géniques (*RET/PTC1*, *RET/PTC3* et *PAX8/PPARgamma*) observés dans certains cas de cancer de la thyroïde [Nikiforov *et al.*, 2011; Nikiforov *et al.*, 2009]. Malgré une spécificité élevée (de 97 à 99 %), la sensibilité du test dans les nodules à cytologie indéterminée était plutôt limitée (de 57 à 63 %); par conséquent, l'analyse était à ses débuts considérée comme un *rule-in test* (VPP de 87 à 88 %). L'arrivée du séquençage de nouvelle génération a grandement facilité l'évaluation concomitante d'un plus large éventail d'altérations génétiques. La version 1 (v1), qui ciblait 284 mutations dans 12 gènes liés au cancer de la thyroïde, a ensuite évolué pour inclure 13 (v2) ou 14 gènes (v2.1) ainsi que 42 produits de fusion génique avant d'aboutir à la version

présentement disponible (v3) [Steward *et al.*, 2019; Nikiforov *et al.*, 2015; Nikiforov *et al.*, 2014; Nikiforova *et al.*, 2013]. Cette dernière version comprend l'évaluation des séquences cibles de 112 gènes sur une plateforme Ion Proton (Thermo Fisher Scientific) et permet l'identification de 5 classes d'altération génétique : variations nucléotidiques simples, insertions/délétions, fusions géniques, altération du niveau d'expression et du nombre de copies. L'algorithme utilisé pour interpréter le résultat du test selon la version employée est présenté dans le tableau ci-dessous.

Tableau 3 Interprétation du résultat de l'analyse ThyroSeq selon la version employée

ThyroSeq	Nodule moléculairement suspect	Nodule moléculairement bénin
Version 2	- Présence d'une mutation dans un des gènes suivants à une fréquence allélique $\geq 10\%$: <i>BRAF</i>, <i>NRAS</i>, <i>HRAS</i>, <i>KRAS</i>, <i>PTEN</i>, <i>TSHR</i>, <i>RET</i>; ou - Présence d'une mutation dans un des gènes suivants à une fréquence allélique $\geq 5\%$: <i>TP53</i>, <i>AKT1</i>, <i>CTNNB1</i>, <i>PIK3CA</i>, <i>TERT</i>; ou - Couverture d'au moins 50X du point de fusion transcriptionnel de 2 gènes.	- Aucune mutation détectée; ou - Présence d'une mutation dans le gène <i>GNAS</i> (peu importe la fréquence allélique).
Version 2.1	- Présence d'une mutation dans un des gènes suivants à une fréquence allélique $\geq 10\%$: <i>NRAS</i>, <i>HRAS</i>, <i>KRAS</i>, <i>PTEN</i>, <i>TSHR</i>, <i>EIF1AX</i>; ou - Présence d'une mutation dans un des gènes suivants à une fréquence allélique $\geq 5\%$: <i>BRAF</i>, <i>TP53</i>, <i>AKT1</i>, <i>CTNNB1</i>, <i>PIK3CA</i>, <i>RET</i>, <i>TERT</i>; ou - Couverture d'au moins 50X du point de fusion transcriptionnel de 2 gènes.	- Aucune mutation détectée; ou - Présence d'une mutation dans le gène <i>GNAS</i> (peu importe la fréquence allélique).
Version 3*	$(x^{\text{SNV/I}})n + x^{\text{GF}} + x^{\text{GEA}} + x^{\text{CNA}} \geq 2$	$(x^{\text{SNV/I}})n + x^{\text{GF}} + x^{\text{GEA}} + x^{\text{CNA}} < 2$

CNA : copy number alteration; GEA : gene expression alteration; GF : gene fusion; SNV/I : single-nucleotide variant/Indel.

* Une valeur pondérée (x) est associée à chaque altération génétique (SNV/I, GF, GEA, CNA), selon son degré d'association avec la malignité (0 : aucune; 1 : faible; 2 : élevé); « n » correspond au nombre de SNV/I. Le score global représente la somme des valeurs individuelles obtenues pour chaque type d'altération.

Afirma

L'analyse Afirma GEC, conçue par une entreprise privée de biotechnologie (Veracyte Inc., San Francisco, CA), repose sur la mesure du niveau d'expression de certains ARNm au moyen de micropuces à ADN [Alexander *et al.*, 2012; Chudova *et al.*, 2010]. La classification des nodules procède par étapes à l'aide de différentes cassettes d'expression. Les six premières cassettes agissent comme filtre et visent l'identification des carcinomes médullaires de la thyroïde, du tissu parathyroïdien, des cellules de Hürthle et des métastases issues d'un mélanome malin, d'un carcinome rénal ou mammaire. Les échantillons qui demeurent négatifs après ces étapes sont ensuite évalués pour l'expression de 142 gènes et analysés au moyen d'un algorithme exclusif

afin de classer les nodules comme moléculairement bénins ou suspects. Récemment, la version GEC a été remplacée par un nouveau test Afirma (GSC) qui emploie une approche de séquençage ciblé de l'ARN sur la plateforme NextSeq 500 d'Illumina [Patel *et al.*, 2018]. Le module principal, basé sur l'analyse d'un total de 10 196 gènes, est accompagné de 7 autres cassettes : tissu parathyroïdien, carcinome médullaire de la thyroïde, mutation BRAF V600E, transcrits de fusion *RET/PTC*, indice de contenu folliculaire, de cellules de Hürthle et de néoplasme à cellules de Hürthle. Les quatre premières composantes sont en amont du module principal alors que les trois dernières ont été incorporées pour améliorer sa performance.

Selon la plus récente information diffusée sur le [site Web de la compagnie](#), l'évaluation des nodules à cytologie indéterminée par Veracyte inclut désormais une caractérisation (Xpression Atlas) des nodules qui présentent un résultat moléculairement suspect avec l'analyse GSC (*Genomic sequencing classifier*).

ThyGenX/ThyraMIR

Un autre outil disponible pour faciliter le diagnostic des nodules à cytologie indéterminée combine les fonctions des essais ThyGenX et ThyraMIR (Interpace Diagnostics Group Inc., Parsipanny, NJ). ThyGenX (précédemment connu sous le nom miRInform, Asuragen Inc.) évalue la présence de 17 altérations génétiques comprenant 14 variations nucléotidiques simples dans les gènes *BRAF*, *HRAS*, *KRAS* et *NRAS*, et 3 transcrits de fusion (*PAX8/PPARGgamma*, *RET/PTC1* et *RET/PTC3*) [Beaudenon-Huibregtse *et al.*, 2014; Giordano *et al.*, 2014]. Employé seul, ce test possède une sensibilité limitée qui prévient l'usage d'un résultat négatif pour exclure la malignité de façon sécuritaire. L'essai ThyraMIR a donc été introduit dans le but de compléter et d'améliorer la performance diagnostique de ThyGenX [Labourier *et al.*, 2015]. Selon le besoin, les échantillons évalués par ThyGenX sont ensuite analysés pour l'expression de 10 micro-ARN différentiellement exprimés dans certains sous-types de tumeurs thyroïdiennes (miR-29b-1-5p, miR-31-5p, miR-138-1-3p, miR-139-5p, miR-146b-5p, miR-155, miR-204-5p, miR-222-3p, miR-375 et miR-551b-3p). L'algorithme qui permet la classification des nodules en fonction du niveau de risque associé au profil moléculaire n'a pas été divulgué. Selon l'information disponible sur le [site Web de la compagnie](#) et celle présentée dans une publication récente, l'analyse ThyGenX/ThyraMIR aurait maintenant évolué vers une nouvelle version (ThyGeNEXT/ThyraMIR) pour laquelle une étude de validation clinique est actuellement en cours [Ablordeppey *et al.*, 2020].

RosettaGX Reveal

De façon similaire à ThyraMIR, le test RosettaGX Reveal (Rosetta Genomics Ltd., Philadelphie, PA) emploie le profil d'expression des micro-ARN pour la classification des nodules à cytologie indéterminée [Lithwick-Yanai *et al.*, 2017]. D'ailleurs, 6 des 24 micro-ARN considérés dans l'analyse sont aussi inclus dans le panel ThyraMIR. Contrairement aux trois autres tests commercialisés, qui nécessitent la collecte de matériel biopsique supplémentaire comme substrat pour l'évaluation moléculaire, l'analyse RosettaGX Reveal est réalisée en utilisant le matériel cellulaire récupéré directement à partir du

frottis coloré. Ce procédé offre l'avantage de réduire le nombre de prélèvements requis pour chaque patient et permet d'analyser le profil moléculaire à partir de la même population caractérisée lors de l'examen cytologique. Sur le [site Web de Rosetta Genomics](#), on apprend cependant que la compagnie a déclaré faillite en mai 2018 et liquidé ses actifs. Pour cette raison, les données scientifiques publiées relativement à ce test ne seront pas exposées dans le présent document.

Contexte de la demande

Au Québec, des demandes de remboursement ont récemment été déposées auprès de la RAMQ pour l'analyse ThyroSeq (v3). La Direction de la biovigilance et de la biologie médicale (DBBM) de la Direction générale des affaires universitaires, médicales, infirmières et pharmaceutiques (DGAUMIP) a demandé à l'INESSS d'évaluer la pertinence de recourir aux tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée.

Le présent avis a pour objectif de faciliter la décision du demandeur 1) en évaluant la performance des analyses moléculaires disponibles pour prédire le diagnostic histopathologique des nodules bénins (validité clinique et analytique); 2) en évaluant la capacité de ces tests à réduire le nombre de chirurgies inutiles pratiquées; 3) en examinant l'aspect sécuritaire associé à la surveillance active des nodules à cytologie indéterminée qui sont moléculairement bénins; 4) en dressant un portrait de la prise en charge actuelle des nodules thyroïdiens à cytologie indéterminée; et 5) en estimant l'impact économique associé au remboursement de ce type d'analyse.

Il est à noter que certains aspects ont été volontairement exclus de la présente évaluation. Notamment, le rôle des tests moléculaires réalisés sur des patients d'âge pédiatrique n'a pas été examiné, d'abord parce que les analyses disponibles sur le marché ont été validées auprès de cohortes adultes, mais aussi parce que la prise en charge optimale des nodules pédiatriques pourrait différer de celle des nodules adultes qui présentent le même résultat cytologique [Cherella *et al.*, 2019; Francis *et al.*, 2015].

Par ailleurs, l'utilisation des analyses moléculaires dans le but d'identifier les nodules malins n'a pas été abordée. Pour exercer une telle fonction, il est attendu qu'un test devrait avoir une VPP similaire à celle d'un diagnostic cytologique malin (98,6 %) [Haugen *et al.*, 2016; Bongiovanni *et al.*, 2012]. Cependant, la VPP des tests évalués est faible (de 37 à 74 %) : la probabilité de ne pas avoir la maladie en cas de résultat positif au test (taux de faux positifs) varie de 26 à 63 % et suggère qu'un résultat positif ne devrait pas servir pour orienter la prise en charge vers un traitement chirurgical.

De surcroît, l'utilisation des tests moléculaires dans le but de déterminer l'étendue de la chirurgie initiale n'a pas été abordée. L'information présentée sur le site Web des entreprises qui font la promotion des analyses moléculaires évaluées indique que le profil associé à un résultat moléculairement suspect peut être employé pour estimer la probabilité de cancer et le pronostic de la maladie ([Afirma](#) et [ThyroSeq](#)). Cette estimation, issue d'algorithmes aux propriétés confidentielles, oriente ensuite la décision visant à déterminer si une thyroïdectomie totale ou une lobectomie devrait être pratiquée. Bien que la présence de certaines mutations parmi les carcinomes papillaires ait été

associée (à divers degrés) au pronostic de la maladie⁴, leur détection n'est présentement pas recommandée comme pratique courante pour la stratification du risque de récurrence [NCCN, 2019; Haugen *et al.*, 2016]. De plus, la valeur pronostique des anomalies décelées (seules ou en combinaisons) parmi les nodules à cytologie indéterminée au moyen des tests moléculaires évalués est empreinte d'incertitude [Mitchell et Yip, 2019]. Le séquençage à haut débit est une technologie capable de cibler des milliers de sites (*hotspots*) sur l'ADN et l'ARN et d'identifier une multitude de signatures moléculaires distinctes. Alors que certaines altérations génétiques ont un impact fonctionnel sur la protéine encodée, plusieurs autres sont silencieuses et sans conséquences réelles sur l'activité de la protéine et les mécanismes cellulaires. Ainsi, comme le détail de l'algorithme des tests moléculaires n'a pas été dévoilé, les critères qui associent chaque profil identifié à une catégorie de risque de récurrence ne peuvent être évalués. De plus, la performance des analyses moléculaires disponibles pour déterminer l'étendue de la chirurgie n'a pas été validée par une étude clinique. Selon plusieurs auteurs, les données sont présentement insuffisantes pour considérer cette forme d'utilisation [Ali *et al.*, 2019; Mitchell et Yip, 2019; Nishino et Nikiforova, 2018; Sahli *et al.*, 2018; Kloos, 2017]. Steward et ses collaborateurs, qui ont démontré la validité clinique de l'analyse ThyroSeq v3 pour l'identification des nodules bénins, précisent d'ailleurs que des études futures seront requises pour mieux comprendre comment l'information génétique extraite devrait être intégrée avec les données cliniques et échographiques pour offrir des traitements plus personnalisés aux patients dont les nodules présentent des profils moléculaires distincts [Steward *et al.*, 2019]. S'il est vrai que ces analyses ont la capacité de prédire correctement le pronostic d'un nodule selon la signature moléculaire identifiée, les prochaines études devraient montrer une prévalence plus élevée de cancers à risque élevé parmi les nodules suspects soumis à une thyroïdectomie totale (sur la base d'une signature moléculaire de haut risque), en comparaison avec les nodules suspects soumis à une lobectomie (en raison d'une signature moléculaire de faible risque). En l'absence de telles études, l'évaluation de la pertinence de recourir aux tests moléculaires a porté uniquement sur leur capacité de prédire quels nodules à cytologie indéterminée sont histologiquement bénins.

⁴ Pour un résumé, voir la section [B21] *Potential impact of BRAFV600E and other mutations on risk estimates in PTC* dans les recommandations de l'ATA [Haugen *et al.*, 2016].

1. MÉTHODOLOGIE

1.1. Questions d'évaluation

Les questions d'évaluation ont été formulées en tenant compte des éléments du modèle PICO : la population à l'étude, l'intervention, les comparateurs et les résultats d'intérêt (*outcomes*).

Validité clinique et analytique

- 1- Quelle est la performance des tests disponibles relativement à l'identification des nodules bénins?
- 2- Quelle est la performance analytique des tests disponibles?

Utilité clinique

- 3- Les tests disponibles permettent-ils d'éviter les chirurgies diagnostiques?
- 4- L'utilisation des tests moléculaires est-elle sécuritaire?

Données contextuelles

- 5- Quel est le portrait actuel de la prise en charge des nodules thyroïdiens à cytologie indéterminée?
- 6- Est-ce qu'une analyse similaire à celles commercialisées aux États-Unis pourrait être réalisée dans les laboratoires publics du Québec?

Perspective des patients

- 7- Quelle est l'expérience de soins vécue par les patients lors de la procédure diagnostique d'un nodule thyroïdien?
- 8- Quelles sont les préoccupations des patients concernant l'utilisation des tests moléculaires pour la caractérisation des nodules thyroïdiens à cytologie indéterminée?

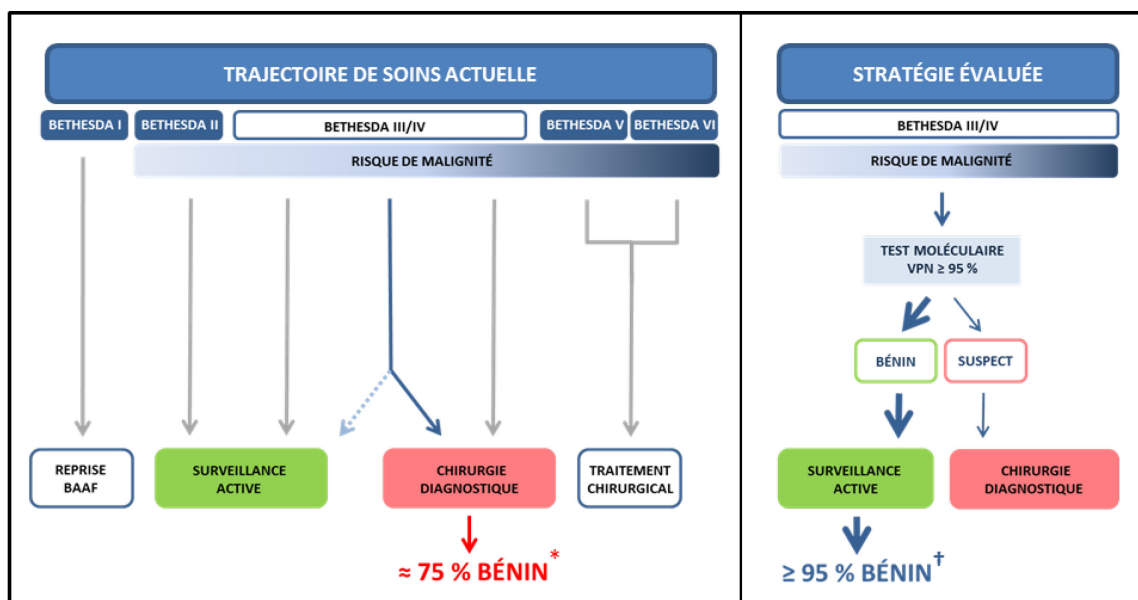
Économie

- 9- L'utilisation des tests moléculaires est-elle efficiente?
- 10- Quel serait l'impact budgétaire associé au remboursement de ce type d'analyse?

1.2. Cadre conceptuel

Le résultat de la biopsie par aspiration à l'aiguille fine, les caractéristiques cliniques et échographiques ainsi que la préférence du patient influent sur la décision de prise en charge des nodules thyroïdiens ([figure B-1](#)). En présence d'un spécimen à cytologie indéterminée (Bethesda III ou IV), les tests moléculaires pourraient compléter l'information disponible sur le risque de malignité et faciliter l'identification des patients qui pourraient éviter la chirurgie diagnostique.

Figure 1 Prise en charge privilégiée des nodules thyroïdiens selon le résultat de la BAAF



BAAF : biopsie par aspiration à l'aiguille fine; VPN : valeur prédictive négative.

* Selon les données de la méta-analyse de Bongiovanni et ses collaborateurs [2012].

† Selon le National Comprehensive Cancer Network (NCCN), les nodules à cytologie indéterminée peuvent faire l'objet d'une surveillance active si l'utilisation d'un test moléculaire donné prédit un risque de malignité ≤ 5 % lorsque le résultat est négatif (VPN ≥ 95 %) [NCCN, 2019].

1.3. Stratégie de repérage de l'information scientifique

La stratégie de recherche de l'information scientifique a été élaborée en collaboration avec un spécialiste en information scientifique. Les mots clés sélectionnés et la stratégie adoptée pour répondre aux questions d'évaluation sont présentés à l'[annexe A](#). Les bases de données bibliographiques MEDLINE (PubMed), Embase et EBM Reviews ont été consultées. Une première recherche, limitée aux articles publiés en français ou en anglais, a été réalisée le 19 mars 2019.

La liste des références des publications retenues a aussi été consultée afin de permettre le repérage des études pertinentes que la stratégie de recherche de la littérature aurait pu omettre.

Une mise à jour de la stratégie de repérage de l'information scientifique a été effectuée périodiquement jusqu'au dépôt de la version préliminaire du document.

Des sources d'information additionnelles ont également été consultées (sites Web des agences, organismes, associations et établissements de santé). Les moteurs de recherche Google et Google Scholar ont été interrogés.

1.4. Critères et processus de sélection des études

Les types de publications examinés comprennent les revues systématiques avec ou sans méta-analyses, les essais cliniques à répartition aléatoire (*randomisés*) ou non et les études observationnelles prospectives et rétrospectives. Les études de cas, les résumés de conférence, les affiches et les revues narratives ont été exclus. La sélection des études pour chaque question a été faite selon les critères PICO.

Tableau 4 Critères d'inclusion et d'exclusion des études scientifiques

Paramètres	Critères d'inclusion	Critères d'exclusion
Population	Patients qui présentent un nodule thyroïdien à cytologie indéterminée (Bethesda III-IV)	Nodules pédiatriques
Intervention	Tests moléculaires (<i>rule-out</i>) : • Afirma • ThyroSeq • ThyGenX/ThyraMIR	<i>Rule-in tests</i> : • ThyroSeq v0 et v1 • ThygenX (miRInform) seul • RosettaGx Reveal
Comparateurs	• Lobectomie diagnostique (validité clinique) • Nodules à cytologie bénigne (Bethesda II) (utilité clinique) • Nodules à cytologie indéterminée (Bethesda III/IV) sans test moléculaire (utilité clinique)	
Résultats d'intérêt (<i>outcome</i>)	Validité clinique : prévalence de la maladie parmi les nodules à cytologie indéterminée, BCR, sensibilité, spécificité, VPP et VPN Validité analytique : sensibilité, spécificité, répétabilité et reproductibilité. Utilité clinique : décision de prise en charge (surveillance vs résection chirurgicale) et taux de faux négatifs Perspective des patients : préférences et qualité de vie Volet économique : impact budgétaire et efficience	Étendue de la chirurgie

BCR : *benign call rate*; VPN : valeur prédictive négative; VPP : valeur prédictive positive.

La sélection des articles a été réalisée par la professionnelle scientifique responsable du dossier (MB) selon les critères présentés précédemment. Un deuxième professionnel scientifique (FB) a validé les raisons qui ont mené à l'exclusion des autres études. Le diagramme de flux des publications sélectionnées ([figure C-1](#)) ainsi que la liste des publications exclues et la raison de leur exclusion ([tableau C-1](#)) sont présentés en annexe.

1.5. Extraction des données

L'extraction des données a été effectuée par la professionnelle scientifique responsable du dossier (MB) selon les paramètres suivants :

Études primaires :

- auteurs, année de publication, plan et objectifs de l'étude;
- caractéristiques de base des patients : nombre de patients, âge, genre;
- caractéristiques des nodules : nombre et taille à l'échographie, résultat de la biopsie selon le système Bethesda, histologie sur spécimen réséqué;
- caractéristiques de l'intervention : type d'analyse moléculaire et critères guidant l'usage (sélectif/systématique);
- résultats d'intérêt : [tableau 4](#);
- conclusions des auteurs.

Lignes directrices et guides de pratique :

- les recommandations avec la méthode de gradation du niveau de preuve et de la force de ces recommandations, s'il y a lieu, ont été présentées.

La vérification des données extraites a été accomplie par un deuxième professionnel scientifique (FB). Lorsque des données pertinentes à l'analyse étaient absentes de la version publiée d'une étude, les auteurs de cette dernière ont été contactés.

1.6. Évaluation de la qualité méthodologique des études

L'évaluation de la qualité méthodologique des études a été réalisée par la professionnelle scientifique responsable du dossier et validée par un deuxième examinateur (FB).

Une version adaptée de l'outil QUADAS-2 a été utilisée pour évaluer la qualité des études de validation clinique [Duh *et al.*, 2017]. La qualité méthodologique des études d'utilité clinique n'a pas été formellement évaluée. Cependant, les principaux critères susceptibles d'affecter la qualité des études retenues ont été identifiés, tels que la durée de l'évaluation (ou du suivi, lorsque disponible), le devis de l'étude, les caractéristiques des nodules examinés, les critères guidant l'usage du test moléculaire (sélectif/systématique) et la présence de conflits d'intérêts. La qualité des études économiques retenues a été évaluée au moyen de l'outil CASP. L'ensemble des résultats de l'évaluation méthodologique sont disponibles à l'[annexe D](#).

1.7. Méthode de contextualisation

Le portrait québécois de la prise en charge actuelle des nodules thyroïdiens a d'abord été documenté par la construction d'un algorithme décisionnel validé par des experts du Québec.

Le nombre annuel de biopsies par aspiration à l'aiguille fine et de chirurgies thyroïdiennes pratiquées entre le 1^{er} janvier 2009 et le 31 décembre 2018 a été obtenu en interrogeant la base de données SMOD de la RAMQ au moyen des codes de rémunération à l'acte employés par les médecins spécialistes : biopsie à l'aiguille de la thyroïde (00237), thyroïdectomie subtotale/bilatérale (06151), hémithyroïdectomie/lobectomie (06152), thyroïdectomie totale (06280), reprise de thyroïdectomie pour goître restant (06281) et thyroïdectomie totale ou subtotale avec résection localisée des ganglions cervicaux (06284). Les cas suivants ont été exclus de l'analyse :

- les actes pratiqués sur des patients de moins de 18 ans (713 actes);
- un même code d'acte attribué plus d'une fois par des dispensateurs différents pour un même patient à une même date (le code attribué par la spécialité du dispensateur le plus approprié a été priorisé);
- les combinaisons d'actes jugées incompatibles lorsque pratiquées chez un même patient (867 actes) :
 - hémithyroïdectomie/lobectomie (06152) ET thyroïdectomie subtotale/bilatérale (06151),
 - hémithyroïdectomie/lobectomie (06152) ET thyroïdectomie totale (06280),
 - hémithyroïdectomie/lobectomie (06152) ET thyroïdectomie avec résection localisée des ganglions cervicaux (06284),
 - thyroïdectomie subtotale/bilatérale (06151) ET thyroïdectomie totale (06280),
 - thyroïdectomie subtotale/bilatérale (06151) ET thyroïdectomie avec résection localisée des ganglions cervicaux (06284),
 - thyroïdectomie totale (06280) ET thyroïdectomie avec résection localisée des ganglions cervicaux (06284),
 - thyroïdectomie totale (06280) ET reprise de thyroïdectomie pour goître restant (06281);
- les actes de reprise de thyroïdectomie pour goître restant (06281) enregistrés entre 2009 et 2018 mais non précédés d'un acte de chirurgie thyroïdienne, ni avant ni après 2009 (346 actes).

1.8. Méthodes d'analyse et de synthèse des données

Les données scientifiques, contextuelles et expérientielles ont été analysées et présentées selon les paramètres des résultats d'intérêt. L'ensemble de la preuve est résumé sous la forme d'une synthèse narrative analytique accompagnée de tableaux.

1.9. Méthodes d'analyse économique

La sélection des études économiques de même que l'extraction et l'analyse des données ont été réalisées par la professionnelle scientifique responsable du dossier (MB). L'évaluation économique primaire et l'analyse d'impact budgétaire ont été effectuées par un professionnel scientifique en économie de la santé (LN).

1.10. Processus de validation scientifique

Une professionnelle scientifique a agi comme rédactrice responsable (MB), appuyée par le professionnel scientifique associé au projet (FB). Des experts du milieu ont été appelés à examiner le document au cours du processus de rédaction. Quatre comités ont été consultés en cours de projet, soit le comité consultatif, le Comité de l'évolution des pratiques en oncologie, le Comité d'excellence clinique en services de santé et le Comité scientifique permanent des analyses de biologie médicale.

Comité consultatif

Un comité consultatif a accompagné l'INESSS dans la réalisation de l'avis avec recommandations afin d'assurer la crédibilité scientifique de ses travaux, la pertinence clinique et de pratique ainsi que l'acceptabilité professionnelle et sociale du produit livré, et ce, en fournissant de l'information, de l'expertise, des opinions ou des perspectives essentielles à la réalisation des travaux.

À cette fin, le comité a été invité à :

- se prononcer sur les thèmes, les questions d'évaluation et les critères d'inclusion et d'exclusion;
- prendre connaissance des résultats de la revue de la littérature;
- fournir de l'information contextuelle et expérientielle.

Comité de l'évolution des pratiques en oncologie (CEPO)

Les membres de ce comité, constitué d'hémo-oncologues, de radio-oncologues, de chirurgiens oncologues et de pharmaciens, ont contribué aux travaux en partageant leur opinion et en soulevant des enjeux cliniques, éthiques et organisationnels.

Comité d'excellence clinique en services de santé (CEC santé)

Le mandat des comités d'excellence clinique consiste à assurer la rigueur scientifique ainsi que l'acceptabilité professionnelle et sociale des productions de l'INESSS. À cette fin, les membres partagent leurs connaissances et leur expertise en appui à l'exercice du cycle d'activités scientifiques de l'Institut en matière de priorisation, production, soutien à l'implantation, mesure et évaluation.

Comité scientifique permanent des analyses de biologie médicale (CSABM)

Le comité scientifique permanent a pour mandat d'analyser les données probantes recueillies par l'INESSS quant à l'utilité clinique, à la validité analytique, aux coûts ainsi qu'aux implications sur les plans organisationnel et éthique.

Il doit également formuler et émettre des recommandations concernant la pertinence d'inscrire les analyses de biologie médicale au Répertoire québécois et système de mesure des procédures de biologie médicale (dans le cas des nouvelles analyses ou des analyses faites hors Québec) ou de les conserver dans ce même répertoire (dans le cas des analyses qui y sont déjà inscrites).

1.11. Processus d'appréciation de la preuve scientifique et formulation des recommandations

L'appréciation de la qualité de la preuve scientifique consiste à juger de la confiance que l'on a relativement au lien possible entre une intervention et les résultats de recherche. Trois critères ont été retenus pour l'évaluation, soit la qualité méthodologique des études, la cohérence des résultats et la transférabilité. Un niveau de preuve scientifique global a ensuite été attribué aux énoncés (élevé, modéré, faible ou insuffisant). Un tableau-synthèse de cet exercice d'appréciation est présenté à l'[annexe G](#).

Des recommandations initiales ont été formulées par les membres de l'équipe de l'INESSS sur la base des données probantes extraites de la littérature scientifique, des opinions et recommandations formulées par d'autres groupes d'experts et des données contextuelles disponibles pour le Québec

L'appréciation de la qualité de la preuve scientifique et la formulation des constats et des recommandations ont été discutées par les membres du Comité d'excellence clinique en services de santé, ceux du CSABM et ceux du CEPO.

1.12. Gestion des conflits d'intérêts

La prévention, la déclaration et la gestion des conflits d'intérêts et de rôles dans le présent dossier font appel à différentes modalités, en accord avec les codes d'éthique applicables, pour assurer l'intégrité des travaux d'évaluation menés et des recommandations formulées, et ainsi préserver la confiance du public envers l'INESSS, ses membres et ses collaborateurs.

Les membres des comités qui ont collaboré aux travaux relatifs à cet avis ont déclaré les intérêts personnels qui pouvaient les placer dans une situation propice au développement de conflits d'intérêts, qu'ils soient commerciaux, financiers, relatifs à la carrière, relationnels ou autres, en plus de déclarer les différentes activités professionnelles ou les rôles qui pouvaient les placer dans une situation propice au développement de conflits de rôles.

Les déclarations remplies par les collaborateurs au dossier ont fait l'objet d'une évaluation par la Direction de l'évaluation et de la pertinence des modes d'intervention en santé de l'INESSS. Les critères pris en considération sont notamment la nature du conflit, la personne concernée par le conflit, le statut temporel, les conséquences potentielles et la perception de l'évaluateur externe.

Les délibérations des groupes de travail en vue de formuler des recommandations ont été structurées à l'aide de processus et de critères clairs, bien définis et transparents, qui permettent d'objectiver les positions prises par leurs membres.

2. RÉSULTATS

2.1. Description des études repérées

La recherche de l'information scientifique a permis de répertorier 365 études, desquelles 46 ont été retenues. L'[annexe C](#) présente le processus de sélection des études sous forme de diagramme de flux ainsi que la liste des études exclues et les raisons de leur exclusion.

2.2. Performance des tests moléculaires commercialisés pour l'identification des nodules bénins à cytologie indéterminée

Idéalement, la performance diagnostique d'un nouveau test devrait être évaluée initialement dans le cadre d'une étude de validation clinique et confirmée ultérieurement par de multiples études de validation indépendantes. Les tests moléculaires qui font l'objet de la présente évaluation ont cependant été rapidement introduits dans la pratique clinique, sans que leur performance soit confirmée par des études indépendantes. Ainsi, la majorité des patients porteurs d'un résultat moléculairement bénin ont été dirigés vers une surveillance active plutôt qu'une chirurgie diagnostique, conformément à la fonction de l'analyse pour le diagnostic des nodules bénins. Par conséquent, la plupart des études publiées en lien avec la performance diagnostique de ces tests ont été menées auprès de patients dont bon nombre de nodules n'avaient pas été réséqués, donc sans confirmation du diagnostic moléculaire. Les données générées dans ce contexte ne sont pas représentatives de la performance du test auprès de la population visée, puisque l'évaluation restreinte aux nodules réséqués est focalisée sur les cas à plus haut risque de malignité [Duh *et al.*, 2017]. L'exclusion des nodules moléculairement bénins non réséqués (histologiquement bénins pour la plupart) entraîne inévitablement une réduction de la spécificité et de la valeur prédictive négative, alors que l'exclusion des nodules moléculairement suspects non réséqués a pour effet de réduire la sensibilité du test. Ainsi, l'évaluation de la performance devrait être réalisée à la suite de la résection de tous les nodules à cytologie indéterminée soumis au test moléculaire, sans égard au résultat obtenu [Kloos, 2017].

Suivant ces considérations, seules les six études de validation clinique originales ont été retenues (trois pour les différentes versions de ThyroSeq, deux pour celles d'Afirma et une pour ThyGenX/ThyraMIR) [Steward *et al.*, 2019; Patel *et al.*, 2018; Labourier *et al.*, 2015; Nikiforov *et al.*, 2015; Nikiforov *et al.*, 2014; Alexander *et al.*, 2012]. Une évaluation de la qualité méthodologique des études sélectionnées est présentée au [tableau D-1](#) en annexe. Les principaux risques de biais identifiés sont liés à la sélection des patients/nodules. Les caractéristiques de performance propres à chacun des tests moléculaires évalués sont résumées au [tableau 5](#).

Les versions 2 et 2.1 de ThyroSeq ont été validées (études monocentriques, rétrospective et prospective) sur des nodules thyroïdiens à cytologie de catégories IV et III, respectivement; aucune information n'était disponible quant à l'âge et au genre des

patients ni à propos de la taille des nodules évalués. Les études de validation concernant ThyroSeq v3 et Afirma GEC étaient de nature prospective et multicentriques; celles réalisées pour ThyGenX/ThyraMIR et Afirma GSC étaient rétrospectives et multicentriques. La majorité des patients inclus étaient du genre féminin (de 70 à 80 %) et tous étaient âgés de plus de 18 ans (âge moyen de chaque cohorte compris entre 50 et 60 ans). La taille médiane des nodules à l'échographie était de 2,1 cm pour ThyroSeq v3 (0,5-7 cm), 2,5 cm pour Afirma GEC (1,0-9,1 cm) et 2,6 cm pour Afirma GSC (1,0-9,1 cm). La proportion de nodules des catégories III et IV était comparable d'une étude à l'autre, mais l'étude de Labourier et ses collaborateurs (ThyGenX/ThyraMIR) comportait une plus faible proportion de nodules de catégorie III (53 %) en comparaison avec les autres études (62 %, 61 % et 60 % pour ThyroSeq v3, Afirma GEC et GSC, respectivement).

2.2.1. Proportion de résultats moléculairement bénins (BCR)

La proportion de nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test est d'un grand intérêt, puisqu'elle représente la population cible qui pourrait éviter une chirurgie diagnostique. En l'absence de faux négatifs (sensibilité 100 %), cette valeur correspond à la prévalence des nodules véritablement bénins. Lorsque l'ensemble des nodules de catégories III et IV a été considéré, les tests ThyroSeq v3 et ThyGenX/ThyraMIR ont produit le plus haut taux de résultats moléculairement bénins, soit 61 %. Dans le cas de ThyroSeq v3, la proportion de résultats bénins était plus élevée pour les nodules de catégorie III que pour ceux de catégorie IV (68 % vs 49 %), alors que ce taux a semblé moins affecté par la catégorie cytologique lorsque les tests Afirma ont été employés (GEC : 43 vs 40 % et GSC : 55 vs 53 %).

2.2.2. Sensibilité et spécificité

La sensibilité d'un test correspond à sa capacité de produire un résultat positif en présence de la maladie (nodules malins et NIFTP). La valeur la plus élevée a été observée avec le test ThyroSeq v3, soit 94 %. Cela signifie qu'en cas de maladie la probabilité de faux négatifs est évaluée à 6 %. La sensibilité des autres tests s'est révélée légèrement inférieure : 89 %, 90 % et 91 % pour ThyGenX/ThyraMIR, Afirma GEC et GSC, respectivement. Par ailleurs ThyGenX/ThyraMIR a obtenu la meilleure spécificité; le test a permis l'identification de 85 % des nodules bénins. En l'absence de la maladie, on peut donc s'attendre à une probabilité de faux positif de 15 % avec ce test. En comparaison, ThyroSeq v3, Afirma GSC et GEC ont montré une spécificité de 82 %, 68 % et 52 %, respectivement.

2.2.3. Valeurs prédictives

Les valeurs prédictives négatives et positives quant à elles renseignent sur la probabilité d'un nodule bénin ou malin/NIFTP lorsque le résultat est négatif ou positif, respectivement. De façon générale, ces valeurs sont tributaires de la sensibilité et de la spécificité du test ainsi que de la [prévalence de la maladie](#) dans la catégorie

indéterminée. Parmi les tests à l'étude, les valeurs prédictives négatives ont varié entre 94 et 97 %, ThyroSeq v3 affichant la meilleure performance; le risque résiduel de maladie en présence d'un résultat négatif (taux de faux négatifs) a été de 3 %. Par ailleurs, des écarts importants ont été observés entre les valeurs prédictives positives des différents tests moléculaires : 74 %, 66 %, 47 % et 37 % pour ThyGenX/ThyraMIR, ThyroSeq v3, Afirma GSC et GEC, respectivement. Ces données traduisent une probabilité somme toute élevée de ne pas avoir la maladie en cas de résultat positif au test : taux de faux positifs égal à 26 %, 34 %, 53 % et 63 %, respectivement.

2.2.4. Prévalence de la maladie parmi les nodules à cytologie indéterminée

Le besoin clinique à combler par les analyses moléculaires évaluées consiste à exclure, de façon sécuritaire, la malignité/NIFTP parmi les nodules thyroïdiens à cytologie indéterminée afin d'identifier correctement ceux qui ne nécessitent pas de résection chirurgicale. La prévalence de la maladie (nodules malins et NIFTP) parmi les nodules de catégories Bethesda III et IV est un facteur déterminant à considérer lors de l'évaluation de la performance des tests moléculaires [Ferris *et al.*, 2015; Marti *et al.*, 2015]. Une augmentation de la proportion de nodules malins/NIFTP a pour effet de réduire la valeur prédictive négative du test et d'augmenter le risque résiduel de maladie dans le cas d'un résultat moléculairement bénin. Selon le NCCN, les nodules à cytologie indéterminée peuvent faire l'objet d'une surveillance active si l'utilisation du test moléculaire prédit un risque résiduel de maladie ≤ 5 % [NCCN, 2019]. L'ATA précise aussi que le test idéal devrait avoir une valeur prédictive négative comparable à celle d'un diagnostic cytologique bénin (96,3 %), donc un risque résiduel de 3,7 % [Haugen *et al.*, 2016].

Les études qui ont évalué la performance des tests moléculaires parmi les nodules à cytologie indéterminée ont rapporté une prévalence de la maladie de 24 % pour Afirma (GEC et GSC), de 28 % pour ThyroSeq v3 et de 32 % pour ThyGenX/ThyraMIR ([tableau 5](#)). Selon le théorème de Bayes et les critères reconnus par le NCCN et l'ATA, l'utilisation de ThyroSeq v3, Afirma GSC et ThyGenX/ThyraMIR serait jugée sécuritaire (VPN ≥ 95 %) jusqu'à l'atteinte d'une prévalence maximale de la maladie de 40 %, 29 % et 28 %, respectivement. Plus de détails sur l'impact de la prévalence locale de la maladie sur la valeur prédictive négative et le risque résiduel sont présentés en annexe au [tableau E-1](#).

Tableau 5 Performance des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée

Étude de validation clinique originale	Bethesda	Prévalence maladie*	BCR†	Sensibilité (IC 95 %)	Spécificité‡ (IC 95 %)	VPN (IC 95 %)	VPP (IC 95 %)
ThyroSeq v2							
Nikiforov <i>et al.</i> , 2014 Rétrospective, monocentrique	143 nodules IV : 143 (100 %)	27 %	71 %	90 % (80-99)	93 % (88-98)	96 % (92-100)	83 % (72-95)
ThyroSeq v2.1							
Nikiforov <i>et al.</i> , 2015 Prospective, monocentrique	96 nodules III : 96 (100 %)	23 %	73 %	91 % (79-100)	92 % (86-98)	97 % (79-100)	77 % (61-93)
ThyroSeq v3							
Steward <i>et al.</i> , 2019 Prospective, multicentrique	247 nodules III : 154 (62 %) IV : 93 (38 %)	28 % III : 23 % IV : 35 %	61 % III : 68 % IV : 49 %	94 % (86-98) III : 91 % (77-97) IV : 97 % (85-100)	82 % (75-87) III : 85 % (77-90) IV : 75 % (63-84)	97 % (93-99) III : 97 % (92-99) IV : 98 % (89-100)	66 % (56-75) III : 64 % (50-77) IV : 68 % (54-80)
Afirma GEC							
Alexander <i>et al.</i> , 2012 Prospective, multicentrique§	210 nodules III : 129 (61 %) IV : 81 (39 %)	24 % III : 24 % IV : 25 %	41 % III : 43 % IV : 40 %	90 % (79-97) III : 90 % (74-98) IV : 90 % (68-99)	52 % (44-60) III : 53 % (43-63) IV : 49 % (36-62)	94 % (88-97) III : 95 % (85-99) IV : 94 % (79-99)	37 % (33-42) III : 38 % (27-50) IV : 37 % (23-52)
Afirma GSC							
Patel <i>et al.</i> , 2018 Rétrospectif, multicentrique¶	190 nodules III : 114 (60 %) IV : 76 (40 %)	24 % III : 25 % IV : 22 %	54 % III : 55 % IV : 53 %	91 % (79-98) III : 93 % (76-99) IV : 88 % (64-99)	68 % (60-76) III : 71 % (60-80) IV : 64 % (51-76)	96 % (90-99) III : 97 % (89-100) IV : 95 % (83-99)	47 % (36-58) III : 51 % (37-65) IV : 42 % (26-59)
ThyGenX/ThyraMIR							
Labourier <i>et al.</i> , 2015 Rétrospective, multicentrique	109 nodules III : 58 (53 %) IV : 51 (47 %)	32 %	61 %	89 % (73-97) III : 94 % (73-100) IV : 82 % (57-96)	85 % (75-92) III : 80 % (64-91) IV : 91 % (76-98)	94 % (85-98) III : 97 % (84-100) IV : 91 % (76-98)	74 % (58-86) III : 68 % (46-85) IV : 82 % (57-96)

BCR : *benign call rate*; GEC : *gene expression classifier*; GSC : *genomic sequencing classifier*; VPN : valeur prédictive négative; VPP : valeur prédictive positive.

* Proportion de nodules malins et de néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire (NIFTP).

† Proportion des nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test et pourraient éviter une chirurgie diagnostique.

‡ Proportion des nodules bénins à cytologie indéterminée qui ont obtenu un résultat négatif au test et pourraient éviter une chirurgie diagnostique.

§ Les données combinées pour l'ensemble des nodules à cytologie indéterminée (III/IV) ont été publiées par Patel et ses collaborateurs en 2018.

¶ Même cohorte de patients qui ont participé précédemment à l'évaluation prospective de la validité clinique du test Afirma GEC.

EN BREF

Deux des trois analyses moléculaires présentement disponibles sur le marché américain ont fait l'objet d'une publication des données de validation clinique (ThyroSeq v3 et Afirma GSC). Leur performance n'a pas été confirmée par des études indépendantes.

Le taux de résultats moléculairement bénins a été de 61 % avec ThyroSeq v3 et de 54 % avec Afirma GSC. Des valeurs similaires de sensibilité (> 90 %) et de VPN (> 95 %) ont été observées alors que la spécificité a été de 82 % avec ThyroSeq v3 et de 68 % avec Afirma GSC. La prévalence de la maladie parmi les nodules à cytologie indéterminée évalués avec ThyroSeq v3 et Afirma GSC était de 28 et 24 %, respectivement.

Selon le théorème de Bayes, la chirurgie peut être évitée de façon sécuritaire si la prévalence locale de la maladie est inférieure à 40 % lorsque le résultat est généré avec l'analyse ThyroSeq v3, et inférieure à 29 % lorsque la version Afirma GSC est utilisée.

La performance des tests moléculaires n'a toutefois pas été évaluée sur les mêmes nodules, ce qui rend la comparaison difficile.

2.3. Performance analytique des tests moléculaires disponibles

L'évaluation de la validité analytique a pour objectif de déterminer si un test donné peut maintenir les mêmes caractéristiques de performance clinique (validité clinique démontrée) en présence de conditions variables. Ce processus est réalisé pour garantir la stabilité de la performance en présence de variations expérimentales issues de la collecte d'échantillons et/ou de leur traitement en laboratoire. Deux publications ont été repérées en lien avec la validité analytique de ThyroSeq v3 et Afirma GSC [Hao *et al.*, 2019; Nikiforova *et al.*, 2018]. Les principaux résultats de ces études sont présentés au [tableau E-2](#) de l'annexe E.

L'évaluation de la sensibilité analytique a montré que les résultats obtenus avec Afirma GSC sont robustes quant aux variations du contenu en ARN (5-30 ng) et tolérants vis-à-vis de la dilution du contenu en cellules malignes avec le tissu sain adjacent (limite de détection du signal de malignité égal à 25 %) [Hao *et al.*, 2019]. En termes de spécificité analytique, l'analyse est demeurée fiable en présence d'une contamination en cellules sanguines ou en ADN génomique pouvant atteindre jusqu'à 75 % et 30 % de l'échantillon, respectivement. Enfin, les scores générés ont été reproductibles pour les échantillons répliqués à l'intérieur d'un même passage sur la plateforme de séquençage (*intra-run*) de même qu'entre passages indépendants (*inter-run*).

Des conclusions similaires ont été tirées à la suite de l'évaluation de la performance analytique de ThyroSeq v3 [Nikiforova *et al.*, 2018]. Le test a permis de classer adéquatement tous les nodules malins en présence d'une quantité d'acides nucléiques comprise entre 2,5 et 10 ng. Le contenu minimal en cellules malignes nécessaire au bon fonctionnement de l'essai a été évalué à 12 % et la tolérance maximale de l'analyse à une contamination en cellules sanguines a été de 88 %. Enfin, l'évaluation de la répétabilité (*intra-run*) et de la reproductibilité (*inter-run*) a révélé que les altérations génétiques ont toutes été détectées avec une variation minimale de la fréquence allélique et du niveau d'expression. Les auteurs ont mentionné qu'aucune variation du score global de l'analyse n'a été observée dans les conditions évaluées (données non présentées).

2.4. Évaluation du nombre de chirurgies pratiquées et du risque résiduel de malignité à la suite d'un résultat moléculairement bénin

En principe, l'utilité clinique des tests moléculaires devrait se traduire par une réduction du nombre de chirurgies thyroïdiennes parmi les nodules à cytologie indéterminée, en comparaison avec la pratique actuelle (en l'absence de telles analyses). Pour en faire la démonstration, la sélection des patients doit être adéquate, c'est-à-dire conforme à l'utilisation ciblée. La cohorte des patients soumis à l'évaluation moléculaire devrait inclure tous ceux qui présentent un nodule à cytologie indéterminée et pour qui le résultat du test est susceptible d'influer sur la prise en charge (chirurgie vs surveillance active). Les patients pour lesquels une chirurgie est indiquée, de même que ceux qui ont une préférence marquée pour la chirurgie ou l'observation, ne devraient pas être sélectionnés pour une analyse moléculaire.

Par ailleurs, la réduction attendue du nombre de chirurgies doit être sécuritaire pour les patients qui présentent un nodule moléculairement bénin désormais placé sous surveillance. On évalue généralement cet aspect en calculant le taux de faux négatifs qui résulte de l'utilisation du test. Or, le taux de faux négatifs attendu est lui-même tributaire de la prévalence locale de maladie parmi les nodules à cytologie indéterminée. Dans un contexte où l'introduction des tests moléculaires dans la pratique clinique a contribué à limiter le nombre de nodules réséqués, la valeur de la prévalence locale de la maladie n'est plus accessible et le taux de faux négatifs est incertain. Dans le but de contourner cette limite, les auteurs ont eu recours à deux approches pour estimer le risque résiduel de malignité suivant un résultat moléculairement bénin. La première consiste à calculer la proportion de cas malins/NIFTP parmi les nodules moléculairement bénins réséqués (6^e colonne des tableaux 6, 8 et 9). Cette approche a été abondamment critiquée, puisqu'elle introduit un biais de sélection qui a pour effet de réduire la valeur prédictive négative et d'augmenter le taux de faux négatifs. À l'opposé, la deuxième méthode génère un taux de faux négatifs dit « optimiste » en présumant que tous les nodules moléculairement bénins non réséqués sont aussi histologiquement bénins : le taux FN-ISG (faux négatif – *incomplete surgical group*) correspond au rapport du nombre de

résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules qui ont obtenu un résultat bénin selon le test moléculaire (septième colonne des tableaux 6, 8 et 9). Étant donné la progression lente des cancers de stade précoce, cette approche a pour effet de sous-estimer la prévalence de la maladie parmi les nodules moléculairement bénins non réséqués.

La majorité des études retenues pour l'appréciation de l'utilité clinique ont été réalisées en l'absence d'un groupe comparateur représentant la pratique actuelle (sans analyse moléculaire). Les résultats sont présentés séparément pour chaque analyse, même lorsque plusieurs tests ont été évalués dans le cadre d'une même étude.

2.4.1. Afirma GEC

Le taux de chirurgies observé pour les nodules à cytologie indéterminée soumis à une évaluation moléculaire avec le test Afirma GEC a fait l'objet de plusieurs études dont 17 ont été sélectionnées pour déterminer l'impact du résultat sur la prise en charge du nodule ([tableau D-2](#) en annexe). La majorité des études étaient rétrospectives; deux études avaient un devis prospectif. La durée médiane du suivi était rarement précisée. Dans la plupart des cas (neuf études), l'évaluation moléculaire a été réalisée de façon sélective (à la discrétion du clinicien ou basé sur la décision partagée) alors que pour trois études la prescription du test était systématique lorsque le résultat de la cytologie était indéterminé; les critères guidant l'usage du test n'ont pas été spécifiés pour cinq études. Le nombre de nodules à cytologie indéterminée inclus dans ces études était très variable (de 50 à 486 nodules). La catégorie Bethesda a été spécifiée dans 15 des 17 études et, à l'exception de 2 d'entre-elles, la cytologie des nodules inclus était majoritairement de catégorie III (étendue : de 56 à 100 %; [tableau 6](#)).

La proportion de nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test (BCR) a été $\geq 40\%$ dans la plupart des études, en accord avec les données de l'étude de validation clinique originale dans laquelle le ratio BCR a été de 41 % (à une prévalence de la maladie de 24 %) [Alexander *et al.*, 2012]. Des valeurs inférieures à 40 % ont cependant aussi été observées, bien que certains facteurs connus soient susceptibles d'avoir contribué à cet effet – notamment dans l'étude de Marti (BCR = 26 %) où la prévalence de la maladie rapportée pour le Memorial Sloan Kettering Cancer Center était élevée (de 30 à 38 %) en comparaison avec le centre Mount Sinai Beth Israel (prévalence de 10 à 19 % et BCR = 52 %). L'étude publiée par McIver a par ailleurs mesuré un BCR de 27 % dans un échantillon de 60 nodules dont la majorité (92 %) présentait une cytologie de catégorie IV. Enfin, la sélection de nodules dont le résultat demeure indéterminé à la suite d'une seconde biopsie par aspiration à l'aiguille fine pourrait expliquer le faible BCR enregistré dans les études menées par Villabona (BCR : 35 %) et par Sultan (BCR : 36 %) [Sultan *et al.*, 2020; Villabona *et al.*, 2016].

La proportion de nodules GEC (*gene expression classifier*) bénins qui ont été réséqués durant la période de suivi a varié de 0 à 33 % entre les études. Le risque résiduel de malignité était compris entre 0 et 53 % pour les nodules GEC bénins réséqués et entre 0 et 17 % pour l'ensemble des nodules GEC bénins (réséqués ou non). Le taux de chirurgies parmi tous les nodules qui ont obtenu un résultat (bénin ou malin) à l'évaluation moléculaire a varié de 37 à 67 %.

Lorsque les nodules des 17 études ont été mis en commun ($n = 3\,401$) de manière à considérer l'effet de la taille de l'échantillon sur les paramètres évalués, on a observé un BCR global de 44 %, un taux de nodules GEC bénins réséqués de 14 % ainsi qu'un taux de faux négatifs de 20 % pour les nodules GEC bénins réséqués (43/215 nodules) et de 3 % pour l'ensemble des nodules moléculairement bénins (43/1 495 nodules). Le taux global de chirurgies pratiquées sur les nodules moléculairement bénins ou malins a été de 51 %.

Tableau 6 Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat Afirma GEC bénin

Étude	Cytologie	Résultat du test	Nodules réséqués	Dx histopathologique		Taux FN-ISG*	Taux de chirurgies†
				Bénin	Malin/ NIFTP		
Arosemena <i>et al.</i> , 2020 Rétrospective	126 nodules III : 126 (100 %) IV : 0	Bénin : 46 (37 %)	15 (33 %)	7 (47 %)	8 (53 %)	8/46 (17 %)	85/126 (67 %)
		Suspect : 80 (63 %)	70 (88 %)	39 (56 %)	31 (44 %)		
San Martin <i>et al.</i> , 2020 Rétrospective	178 nodules III : 103 (58 %) IV : 75 (42 %)	Bénin : 73 (41 %)	5 (7 %)	4 (80 %)	1 (20 %)	1/73 (1 %)	85/178 (48 %)
		Malin : 105 (59 %)	80 (76 %)	48 (60 %)	32 (40 %)		
Sultan <i>et al.</i> , 2020 Rétrospective	50 nodules III : 47 (94 %) IV : 3 (6 %)	Bénin : 18 (36 %)	1 (6 %)	1 (100 %)	0	0/18 (0 %)	19/50 (38 %)
		Suspect : 32 (64 %)	18 (56 %)	7 (39 %)	11 (61 %)		
Endo <i>et al.</i> , 2019 Rétrospective	343 nodules III : 228 (66 %) IV : 115 (34 %)	Bénin : 165 (48 %)	27 (16 %)	24 (89 %)	3 (11 %)	3/165 (2 %)	180/343 (52 %)
		Suspect : 178 (52 %)	153 (86 %)	102 (67 %)	51 (33 %)		
Angell <i>et al.</i> , 2019 Rétrospective	486 nodules III : 314 (65 %) IV : 172 (35 %)	Bénin : 233 (48 %)	28 (12 %)	26 (93 %)	2 (7 %)	2/233 (< 1 %)	249/486 (51 %)
		Suspect : 253 (52 %)	221 (87 %)	146 (66 %)	75 (34 %)		
Harrell <i>et al.</i> , 2019 Rétrospective	481 nodules (III/IV)‡	Bénin : 200 (42 %)	61 (30 %)	42 (69 %)	19 (31 %)	19/200 (10 %)	270/481 (56 %)
		Suspect : 281 (58 %)	209 (74 %)	89 (43 %)	120 (57 %)		
Wei <i>et al.</i> , 2019 Rétrospective	194 nodules III : 155 (80 %) IV : 39 (20 %)	Bénin : 88 (45 %)	3 (3 %)§	2 (67 %)	1 (33 %)	1/88 (1 %)	82/194 (42 %)
		Suspect : 106 (55 %)	79 (75 %)	50 (63 %)	29 (37 %)		
Deaver <i>et al.</i> , 2018 Rétrospective	167 nodules III : 125 (75 %) IV : 42 (25 %)	Bénin : 73 (46 %)	5 (7 %)	5 (100 %)	0	0/73 (0 %)	75/159 (47 %)
		Suspect : 86 (54 %)	70 (81 %)	47 (67 %)	23 (33 %)		
Livhits <i>et al.</i> , 2018 Prospective, randomisée	70 nodules III : 63 (90 %) IV : 7 (10 %)	Bénin : 30 (47 %)	3 (10 %)	3 (100 %)	0	0/30 (0 %)	29/64 (45 %)
		Suspect : 34 (53 %)	26 (76 %)	16 (62 %)	10 (38 %)		
Kay-Rivest <i>et al.</i> , 2017 Rétrospective	109 nodules - Mtl III : 84 (77 %) IV : 25 (23 %)	Bénin : 60 (55 %)	0	0	0	0/60 (0 %)	44/109 (40 %)
		Suspect : 49 (45 %)	44 (90 %)	23 (52 %)	21 (48 %)		
	63 nodules - St.J III : 21 (33 %) IV : 42 (67 %)	Bénin : 29 (46 %)	0	0	0	0/29 (0 %)	33/63 (52 %)

Étude	Cytologie	Résultat du test	Nodules réséqués	Dx histopathologique		Taux FN-ISG*	Taux de chirurgies†
				Bénin	Malin/ NIFTP		
Wu <i>et al.</i> , 2016 Rétrospective	245 nodules III : 217 (89 %) IV : 28 (11 %)	Bénin : 113 (46 %)	24 (21 %)	20 (83 %)	4 (17 %)	4/113 (4 %)	128/245 (52 %)
		Suspect : 132 (54 %)	104 (79 %)	45 (43 %)	59 (57 %)		
Villabona <i>et al.</i> , 2016 Rétrospective	100 nodules III : 100 (100 %) IV : 0	Bénin : 35 (35 %)	4 (11 %)	3 (75 %)	1 (25 %)	1/35 (3 %)	60/100 (60 %)
		Suspect : 65 (65 %)	56 (86 %)	10 (18 %)	46 (82 %)		
Samulski <i>et al.</i> , 2016 Rétrospective	277 nodules¶ III : 159 (57 %) IV : 118 (43 %)	Bénin : 144 (52 %)	16 (11 %)	13 (81 %)	3 (19 %)	3/144 (2 %)	121/277 (44 %)
		Suspect : 133 (48 %)	105 (79 %)	64 (61 %)	41 (39 %)		
Sacks <i>et al.</i> , 2016 Rétrospective	129 nodules (III/IV)	Bénin : 48 (40 %)	4 (8 %)	4 (100 %)	0	0/48 (0 %)	58/120 (48 %)
		Suspect : 72 (60 %)	54 (75 %)	36 (67 %)	18 (33 %)		
Chaudhary <i>et al.</i> , 2016 Rétrospective	158 nodules III : 89 (56 %) IV : 69 (44 %)	Bénin : 63 (43 %)	8 (13 %)	8 (100 %)	0	0/63 (0 %)	81/148 (55 %)
		Suspect : 85 (57 %)	73 (86 %)	45 (62 %)	28 (38 %)		
Marti <i>et al.</i> , 2015 Rétrospective	94 nodules - MSK III : 63 (67 %) IV : 31 (33 %)	Bénin : 24 (26 %)	2 (8 %)	2 (100 %)	0	0/24 (0 %)	44/94 (47 %)
		Suspect : 70 (74 %)	42 (60 %)	18 (43 %)	24 (57 %)		
	71 nodules - MSBI III : 40 (56 %) IV : 31 (44 %)	Bénin : 37 (52 %)	5 (14 %)	5 (100 %)	0	0/37 (0 %)	26/71 (37 %)
McIver <i>et al.</i> , 2014 Prospective	60 nodules III : 5 (8 %) IV : 55 (92 %)	Bénin : 16 (27 %)	4 (25 %)	3 (75 %)	1 (25 %)	1/16 (6 %)	36/60 (60 %)
		Suspect : 44 (73 %)	32 (73 %)	27 (84 %)	5 (16 %)		
TOTAL	3 401 nodules	GEC bénin 1 495/3 368 (44 %)	GEC bénin réséqué 215/1 495 (14 %)	GEC/Histo bénin 172/215 (80 %)	Taux FN 43/215 (20 %)	43/1 495 (3 %)	1 705/3 368 (51 %)

Dx : diagnostic; FN : faux négatif; FN-ISG : faux négatif- *incomplete surgical group*; GEC : *gene expression classifier*; MSBI : Mount Sinai Beth Israel; MSK : Memorial Sloan Kettering Cancer Center; Mtl : Hôpital général juif de Montréal; NIFTP : néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire; St.J : Hôpital St-Clare's Mercy de St-John's.

* Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules qui ont obtenu un résultat bénin au test moléculaire, bien que le nombre de nodules bénins réséqués ait été incomplet au moment de l'analyse.

† Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de nodules réséqués sur l'ensemble des nodules qui ont obtenu un résultat au test moléculaire (bénin ou suspect).

‡ Tous les résultats de BAAF ont été interprétés par le service de cytopathologie (Thyroid Cytopathology Partners; TCP, Texas) désigné par Veracyte Inc.

§ Le nombre de nodules GEC bénins réséqués a été obtenu de l'auteur de correspondance.

¶ Les données spécifiques aux nodules à cytologie indéterminée ont été obtenues de la méta-analyse de Valderrabano et ses collaborateurs [2019].

|| Valeurs calculées par l'INESSS.

2.4.2. Impact de l'introduction du test Afirma GEC

Dans le but de mesurer l'impact de l'introduction du test Afirma GEC, trois des études retenues ont aussi comparé les résultats des patients soumis ou non à une évaluation moléculaire ([tableau 7](#)). Considérant le rôle de l'analyse pour l'identification des nodules bénins, il est attendu qu'un effet favorable se traduirait par une diminution du taux de chirurgies thyroïdiennes accompagnée d'une augmentation de la proportion de cas malins parmi les nodules réséqués (résultant d'une meilleure sélection des patients à opérer).

L'étude menée par Sacks et ses collaborateurs incluait 567 nodules à cytologie indéterminée (sur 4 292 biopsies; 13,2 %) évalués au cours de la période pré (01/2012-07/2013) ou post-introduction du test (07/2013-01/2014) Afirma [Sacks *et al.*, 2016]. À la suite de l'implémentation, moins de la moitié (45 %) des nodules à cytologie indéterminée ont été soumis au test moléculaire. Le suivi médian après la BAAF a été de 20,9 mois (IC 95 % 19,7-22,1). Quatre-vingt-quatorze patients ont été perdus de vue au suivi après la BAAF et exclus de l'analyse. Aucune différence significative du taux de chirurgies ou de la proportion de cas malins n'a été détectée entre les cohortes post (n = 253) et pré-introduction du test (n = 220) Afirma, et ce, même lorsque les résultats ont été analysés par catégories cytologiques. Des données similaires ont été obtenues lorsque le sous-groupe de patients composé uniquement de nodules soumis au test Afirma GEC (n = 129) a été comparé aux groupes sans évaluation moléculaire de la période pré (n = 220) ou post-introduction (n = 124) du test Afirma.

Au cours de la période post-Afirma (07/2012-06/2015) de leur centre, Chaudhary et ses collaborateurs ont identifié 376 nodules à cytologie indéterminée sur un total de 2 560 biopsies (15 %) [Chaudhary *et al.*, 2016]. Cent-cinquante-huit (42 %) de ces nodules ont été évalués avec Afirma GEC. En comparaison avec les nodules à cytologie indéterminée (n = 280) de la période pré-Afirma du même établissement (07/2009-06/2012), aucune différence significative n'a été rapportée concernant le taux de chirurgies pratiquées ou la proportion de nodules malins réséqués à la suite de l'introduction du test. Une réduction significative du taux de résections chirurgicales a cependant été observée dans le sous-groupe de nodules à cytologie de catégorie IV (76 vs 52 %; p = 0,001), sans toutefois affecter le taux de malignité des nodules réséqués.

Enfin, Yang et ses collaborateurs ont identifié 189 nodules à cytologie indéterminée soumis à l'évaluation moléculaire parmi les 1 693 biopsies réalisées au cours de la période qui a suivi l'introduction de l'analyse GEC (08/2012-04/2014) dans leur établissement [Yang *et al.*, 2016]. Les données de ce groupe de nodules ont été comparées à celles colligées pour 502 nodules à cytologie indéterminée de la période pré-Afirma (2003-2012) [Sullivan *et al.*, 2014]. Les résultats ont été présentés par sous-groupes cytologiques. Les auteurs ont mentionné avoir observé une augmentation significative de la proportion de nodules malins réséqués après l'implémentation de l'analyse GEC, qu'ils ont expliquée par la réduction du taux de chirurgies durant la même période. Cependant, la valeur p a été calculée en incluant les deux paramètres évalués lors des périodes pré et post-Afirma. Lorsque ces paramètres ont été évalués séparément

(calcul interne), la différence entre les taux de chirurgies des deux périodes est demeurée significative pour les nodules à cytologie de catégorie III ($p = 0,0010$), mais pas pour ceux de catégorie IV ($p = 0,2657$), et la proportion de nodules malins réséqués est restée similaire (III : $p = 0,1077$ et IV : $p = 0,3377$).

En résumé, les résultats présentés au tableau 7 n'ont pas permis de démontrer une amélioration significative du taux de chirurgies ou de la proportion de cas malins réséqués suivant l'introduction de l'analyse Afirma GEC. Par ailleurs, plusieurs facteurs ont limité la qualité des études qui ont évalué l'impact de l'analyse. La durée du suivi médian et les caractéristiques de base des nodules n'ont pas fait l'objet d'une comparaison entre les cohortes étudiées. Aussi, les raisons qui ont guidé la décision de procéder ou non à l'analyse moléculaire lors de la période post-Afirma n'ont pas été précisées. Les études qui ont exclu de la période post-Afirma les nodules sans évaluation moléculaire s'exposent ainsi à un risque de biais, puisqu'à l'opposé ces patients étaient représentés dans la cohorte pré-Afirma. Pour ces raisons, les cohortes comparées étaient potentiellement hétérogènes selon plusieurs aspects.

Tableau 7 Impact de l'analyse Afirma GEC sur le taux de chirurgies thyroïdiennes et la prévalence de cas malins chez les patients présentant un nodule à cytologie indéterminée

Étude	Paramètres évalués	PRÉ Afirma GEC (A)	POST Afirma GEC			p		
			SANS évaluation moléculaire (B)	AVEC évaluation moléculaire (C)	TOTAL (D)	C vs A	C vs B	D vs A
Sacks <i>et al.</i> , 2016 Rétrospective	Bethesda III/IV							
	Chx thyroïdienne	83/220 (37,7 %)	54/124 (44 %)	60/129* (46,5 %)	114/253 (45,1 %)	0,12	0,70	0,11
	Malignité	21/83 (25,3 %)	23/54 (43 %)	18/60 (30,0 %)	41/114 (36,0 %)	0,57	0,18	0,12
Chaudhary <i>et al.</i> , 2016 Rétrospective	Bethesda III/IV							
	Chx thyroïdienne	171/280 (61 %)	-	86/158 [†] (54 %)	-	n.s.	-	-
	Malignité	55/171 (32 %)	-	29/86 (34 %)	-	n.s.	-	-
Yang <i>et al.</i> , 2016 Rétrospective [‡]	Bethesda III							
	Chx thyroïdienne	156/318 (49 %)	-	55/165 (33 %)	-	0,04[§]	-	-
	Malignité	54/156 (35 %)	-	26/55 (47 %)	-		-	-
	Bethesda IV							
	Chx thyroïdienne	116/184 (63 %)	-	12/24 (50 %)	-	0,03[§]	-	-
	Malignité	38/116 (33 %)	-	6/12 (50 %)	-		-	-

Chx : chirurgie; GEC : *gene expression classifier*; n.d. : non déterminé; n.s. : non significatif.

* Incluant 9 nodules qui ont obtenu un résultat *non-diagnostic* au test moléculaire, dont 2 ont été réséqués; les taux de chirurgies en fonction du résultat au test sont présentés au [tableau 6](#).

† Incluant 10 nodules qui ont obtenu un résultat *non-diagnostic* au test moléculaire dont 5 ont été réséqués; les taux de chirurgies en fonction du résultat au test sont présentés au tableau 6.

‡ Les données de la période post-Afirma incluent 15 nodules (13 Bethesda III et 2 Bethesda IV) qui ont obtenu un résultat *non-diagnostic* au test moléculaire et dont aucun n'a été réséqué; le taux de chirurgies pour les nodules GEC bénins et malins a été de 6 et 66 %, respectivement. La cohorte de patients du groupe GEC de cette étude chevauche celle décrite par Wu et ses collaborateurs [2016] dont les données sont présentées au tableau 6.

§ Les valeurs en pourcentage des deux paramètres évalués lors des périodes pré et post-Afirma ont été employées pour la méthode exacte de Fisher.

2.4.3. Afirma GSC

La gestion chirurgicale des nodules à cytologie indéterminée à la suite de l'emploi de la plus récente version de l'analyse Afirma (GSC) a été évaluée rétrospectivement dans cinq établissements américains [San Martin *et al.*, 2020; Angell *et al.*, 2019; Endo *et al.*, 2019; Harrell *et al.*, 2019; Wei *et al.*, 2019]. Le nombre de nodules inclus dans chaque étude a varié de 78 à 164. Une seule étude a spécifié avoir fait un usage sélectif du test basé sur la décision partagée, en l'absence d'indication claire pour une chirurgie; les critères qui ont guidé l'usage du test étaient inconnus ou absents (usage systématique) dans les autres études ([tableau D-2](#) en annexe). La proportion de nodules à cytologie de catégorie III a varié de 59 à 85 %; elle n'a pas été spécifiée pour une des cinq études.

La proportion de nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test (BCR) a été supérieure à 60 % dans les 5 études retenues (étendue : de 61 à 76 %; [tableau 8](#)); elle était de 54 % dans l'étude de validation clinique originale (à une prévalence de la maladie de 24 %) [Patel *et al.*, 2018]. La proportion de nodules GSC bénins réséqués durant la période de suivi (durée médiane non spécifiée) était comprise entre 2 et 10 %. Le risque résiduel de malignité a varié de 0 à 37 % pour les nodules GSC bénins réséqués; il était inférieur à 5 % pour l'ensemble des nodules GSC bénins (réséqués ou non). Rappelons cependant qu'un nombre limité de nodules moléculairement bénins ont été réséqués dans le cadre de ces études (26 nodules au total). Le taux de chirurgies pratiquées sur tous les nodules soumis à l'analyse GSC a varié de 18 à 35 %.

Les données obtenues avec Afirma GSC ont aussi été comparées à celles générées avec Afirma GEC ([tableau 6](#)). Une augmentation significative du ratio BCR avec la version GSC a été rapportée par les cinq études. Celle-ci était accompagnée d'une réduction du taux de chirurgies pratiquées pour les nodules soumis à l'analyse GSC (moléculairement bénins et malins) en comparaison avec la version GEC; selon les études, la réduction⁵ observée a varié de 27 à 66 %. Certains éléments sont cependant de nature à limiter la portée de ces comparaisons. Étant donné la mise en application récente de l'analyse GSC, la période d'évaluation a été plus courte (2017-2018) que pour l'analyse GEC (2011-2017) et le nombre de nodules inclus plus restreint. Par conséquent, la proportion de nodules GSC bénins réséqués est susceptible d'augmenter en réponse au développement éventuel de facteurs de risque. Aussi, à l'exception de l'étude menée par Angell et ses collaborateurs, dans laquelle une distribution similaire des caractéristiques de base des patients / nodules entre les deux groupes a été observée, les autres études considérées ont soit négligé cet aspect de l'évaluation, soit réalisé une comparaison très sommaire des cohortes. L'absence de démonstration d'une répartition semblable de facteurs tels que l'âge, le genre, le résultat de la cytologie, la taille du nodule et les caractéristiques échographiques entre les groupes comparés ne permet donc pas d'exclure la possibilité d'un déséquilibre capable d'influer sur les résultats présentés.

⁵ Réduction relative du taux de chirurgie : $([\text{taux de chirurgies avec GEC}] - [\text{taux de chirurgies avec GSC}]) / [\text{taux de chirurgies avec GEC}]$.

Tableau 8 Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat Afirma GSC bénin

Étude	Cytologie	Résultat du test	Nodules réséqués	Dx histopathologique		Taux FN-ISG*	Taux de chirurgies†
				Bénin	Malin/NIFTP		
San Martin <i>et al.</i> , 2020 Rétrospective	121 nodules III : 76 (63 %) IV : 45 (37 %)	Bénin : 82 (68 %)	8 (10 %)	5 (63 %)	3 (37 %)	3/82 (4 %)	42/121 (35 %)
Endo <i>et al.</i> , 2019 Rétrospective	164 nodules III : 124 (76 %) IV : 40 (24 %)	Bénin : 125 (76 %) Suspect : 39 (24 %)	4 (3 %) 25 (64 %)	4 (100 %) 10 (40 %)	0 15 (60 %)	0/125 (0 %)	29/164 (18 %)
Angell <i>et al.</i> , 2019 Rétrospective	114 nodules III : 67 (59 %) IV : 47 (41 %)	Bénin : 75 (66 %) Suspect : 39 (34 %)	5 (7 %) 32 (82 %)	4 (80 %) 16 (50 %)	1 (20 %) 16 (50 %)	1/75 (1 %)	37/114 (32 %)
Harrell <i>et al.</i> , 2019 Rétrospective	139 nodules (III/IV)‡	Bénin : 85 (61 %) Suspect : 54 (39 %)	8 (9 %) 37 (69 %)	7 (88 %) 9 (24 %)	1 (12 %) 28 (76 %)	1/85 (1 %)	45/139 (32 %)
Wei <i>et al.</i> , 2019 Rétrospective	78 nodules III : 66 (85 %) IV : 12 (15 %)	Bénin : 52 (67 %) Suspect : 26 (33 %)	1 (2 %)§ 14 (54 %)	1 (100 %) 6 (43 %)	0 8 (57 %)	0/52 (0 %)	15/78 (19 %)
TOTAL¶	616 nodules	GSC bénin 419/616 (68 %)	GSC bénin réséqué 26/419 (6 %)	GSC/Histo bénin 21/26 (81 %)	Taux FN 5/26 (19 %)	5/419 (1 %)	168/616 (27 %)

Dx : diagnostic; FN : faux négatif; FN-ISG : faux négatif- *incomplete surgical group*; GSC : *genomic sequencing classifier*; NIFTP : néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire.

* Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules qui ont obtenu un résultat bénin au test moléculaire, bien que le nombre de nodules bénins réséqués ait été incomplet au moment de l'analyse.

† Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de nodules réséqués sur l'ensemble des nodules qui ont obtenu un résultat au test moléculaire (bénin ou suspect).

‡ Tous les résultats des BAAF ont été interprétés par le service de cytopathologie (Thyroid Cytopathology Partners; TCP, Texas) désigné par Veracyte Inc.

§ Le nombre de nodules GSC bénins réséqués a été obtenu de l'auteur de correspondance.

¶ Valeurs calculées par l'INESSS.

2.4.4. ThyroSeq v2.1

Deux études rétrospectives ont été retenues pour l'évaluation de l'utilité clinique de l'analyse ThyroSeq v2.1 [Taye *et al.*, 2018; Valderrabano *et al.*, 2017]. L'emploi du test était systématique dans l'étude menée par Valderrabano, mais variable (sélectif ou systématique) selon le médecin prescripteur dans celle de Taye et ses collaborateurs ([tableau D-2](#) en annexe). Les nodules à cytologie indéterminée inclus dans chacune de ces études ont été au nombre de 156 et 190, dont 79 et 55 % présentaient une cytologie de catégorie III, respectivement. Rappelons que la version 2.1 a été validée dans une cohorte de patients qui présentaient exclusivement des nodules à cytologie de catégorie III [Nikiforov *et al.*, 2015].

Les nodules à cytologie indéterminée qui ont obtenu un résultat négatif/bénin au test moléculaire (BCR) ont été représentés dans une proportion égale à 67 et 76 %, respectivement ([tableau 9](#)); le ratio BCR de l'étude de validation clinique originale a été de 73 % (à une prévalence de la maladie de 23 %) [Nikiforov *et al.*, 2015]. La proportion de nodules ThyroSeq bénins réséqués au cours de la période de suivi (durée médiane non spécifiée) a été de 23 et 48 %. Dans les 2 cas, le risque résiduel de malignité a été de 9 % pour les nodules moléculairement bénins réséqués et de 2 et 4 % pour tous les nodules ThyroSeq bénins (peu importe le statut chirurgical). Le taux de chirurgies observé parmi les nodules ThyroSeq bénins et malins a été de 39 et 54 %.

Par ailleurs, une étude prospective *randomisée* a été réalisée dans le but de déterminer laquelle des analyses, Afirma GEC ou ThyroSeq v2.1, permettait à un plus grand nombre de patients d'éviter une chirurgie inutile [Livhits *et al.*, 2018]. L'analyse moléculaire a été prescrite de façon systématique lorsque le résultat de la cytologie était indéterminé ([tableau D-2](#) en annexe). Les données des 70 nodules inclus dans le groupe GEC ont été présentées dans le tableau 6. Les caractéristiques de base des patients étaient similaires dans les deux groupes examinés.

Soixante-dix-neuf nodules ont été inclus dans le groupe ThyroSeq v2.1, dont la majorité (90 %) présentait une cytologie de catégorie III ([tableau 9](#)). Le ratio BCR observé a été de 80 % (vs 47 % pour GEC) et la proportion de nodules ThyroSeq bénins réséqués au cours de la période de suivi (durée médiane non spécifiée) a été de 15 % (vs 10 % pour les nodules GEC bénins). Les données du suivi échographique biannuel étaient disponibles pour 46 nodules moléculairement (ThyroSeq ou GEC) bénins : 8 nodules ont eu un suivi de 6 mois, 27 nodules ont eu un suivi de 12 mois et 11 nodules un suivi de 18 mois. Parmi ceux-ci, un seul (ThyroSeq bénin) a présenté une croissance significative à 12 mois, ce qui a motivé la reprise de la biopsie et de l'évaluation moléculaire (Bethesda III; GEC bénin); le nodule est cependant demeuré en observation. Au total, peu de nodules bénins ont été réséqués (9 pour ThyroSeq et 3 pour GEC); le risque résiduel de malignité était nul dans les deux cas. Globalement, 30 % des nodules ThyroSeq bénins ou malins ont été réséqués (45 % pour GEC).

Tableau 9 Évaluation du taux de chirurgies et du risque résiduel de malignité à la suite d'un résultat ThyroSeq v2.1 bénin

Étude	Cytologie	Résultat du test	Nodules réséqués	Dx histopathologique		Taux FN-ISG*	Taux de chirurgies†
				Bénin	Malin/ NIFTP		
Livhits <i>et al.</i> , 2018 Prospective, <i>randomisée</i>	79 nodules III : 71 (90 %) IV : 8 (10 %)	Bénin : 61 (80 %)	9 (15 %)	9 (100 %)	0	0/61 (0 %)	23/76 (30 %)
		Suspect : 15 (20 %)	14 (93 %)	6 (43 %)	8 (57 %)		
Taye <i>et al.</i> , 2018 Rétrospective, multicentrique	156 nodules III : 124 (79 %) IV : 32 (21 %)	Bénin : 102 (67 %)	23 (23 %)	21 (91 %)	2 (9 %)	2/102 (2 %)	60/153 (39 %)
		Suspect : 51 (33 %)	37 (73 %)	27 (73 %)	10 (27 %)		
Valderrabano <i>et al.</i> , 2017 Rétrospective	190 nodules III : 104 (55 %) IV : 86 (45 %)	Bénin : 145 (76 %)	69 (48 %)	63 (91 %)	6 (9 %)	6/145 (4 %)	102/190 (54 %)
		Suspect : 45 (24 %)	33 (73 %)	19 (58 %)	14 (42 %)		
TOTAL‡	425 nodules	ThyroSeq bénin 308/419 (74 %)	ThyroSeq bénin réséqué 101/308 (33 %)	Thyro/Histo bénin 93/101 (92 %)	Taux FN 8/101 (8 %)	8/308 (3 %)	185/419 (44 %)

Dx : diagnostic; FN : faux négatif; FN-ISG : faux négatif- *incomplete surgical group*; NIFTP : néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire.

* Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de résultats faussement négatifs après l'évaluation des spécimens réséqués sur l'ensemble des nodules qui ont obtenu un résultat bénin au test moléculaire, bien que le nombre de nodules bénins réséqués ait été incomplet au moment de l'analyse.

† Valeurs calculées par l'INESSS lorsque non rapportées par les auteurs; rapport du nombre de nodules réséqués sur l'ensemble des nodules qui ont obtenu un résultat au test moléculaire (bénin ou suspect).

‡ Valeurs calculées par l'INESSS.

2.4.5. ThyroSeq v3

Une étude québécoise portant sur 50 nodules à cytologie indéterminée évalués entre janvier et février 2019 avec la plus récente version de l'analyse ThyroSeq (v3) a été repérée [Chen *et al.*, 2020]. L'emploi du test était restreint aux nodules qui présentaient des caractéristiques échographiques associées à un risque de malignité modéré ou élevé (TI-RADS 4 et 5). La taille moyenne des nodules était de 2,5 cm (étendue : de 1,1 à 4,2 cm) et la proportion de nodules à cytologie de catégorie Bethesda III était de 42 %.

Après l'exclusion de 2 nodules pour lesquels l'évaluation moléculaire a échoué, le taux de résultats moléculairement bénins a été de 58 % (28/48); celui-ci était de 61 % dans l'étude de validation clinique originale [Steward *et al.*, 2019]. La distribution des résultats moléculaires selon la catégorie de risque échographique (TI-RADS) n'a pas été présentée. La proportion de nodules moléculairement bénins réséqués durant la période de suivi (durée médiane non spécifiée) a été de 7 % (2/28); les 2 nodules réséqués étaient histologiquement bénins. Par ailleurs, 19 des 20 nodules moléculairement suspects ont été réséqués; tous étaient histologiquement malins/NIFTP. Le taux de chirurgies pratiquées sur les nodules soumis à l'analyse ThyroSeq v3 a été de 44 % (21/48). Un biais lié à la sélection des cas à risque élevé est possiblement responsable du taux élevé de chirurgies rapporté dans cette étude. Le risque résiduel de maladie associé aux nodules moléculairement bénins qui n'ont pas été réséqués n'a pas été caractérisé dans ce contexte.

EN BREF

Le nombre d'études qui ont permis d'évaluer l'utilité clinique des tests moléculaires est réparti comme suit : 17 études pour Afirma GEC, 5 études pour Afirma GSC, 3 études pour ThyroSeq v2.1 et 1 étude pour ThyroSeq v3.

Le taux de résultats moléculairement bénins calculé pour l'ensemble des nodules évalués avec chaque test a été comparable ou supérieur à celui rapporté pour la cohorte de validation initiale : 44 % pour Afirma GEC (vs 41 %), 68 % pour Afirma GSC (vs 54 %), 74 % pour ThyroSeq v2.1 (vs 73 %; catégorie III seulement) et 58 % pour ThyroSeq v3 (vs 61 %).

Le taux historique de nodules à cytologie indéterminée réséqués a été évalué à 55 % (méta-analyse de Bongiovanni [2012]).

Le taux de chirurgies pratiquées pour l'ensemble des nodules évalués avec chaque test a été de 51 % pour Afirma GEC, de 27 % pour Afirma GSC et de 44 % pour ThyroSeq v2.1 et v3.

Le taux de nodules moléculairement bénins réséqués a été de 14 % pour Afirma GEC, de 6 % pour Afirma GSC, de 33 % pour ThyroSeq v2.1 et de 7 % pour ThyroSeq v3.

Le risque résiduel de maladie associé aux nodules moléculairement bénins non réséqués est incertain.

2.5. Comparaison des nodules à cytologie indéterminée moléculairement bénins avec les nodules à cytologie bénigne

Une autre façon d'aborder la question de l'utilité clinique des tests moléculaires implique une comparaison des données obtenues pour les nodules à cytologie indéterminée moléculairement bénins avec celles issues des nodules à cytologie bénigne.

Les principaux enjeux de la prise en charge des nodules à cytologie indéterminée moléculairement bénins concernent la durée du suivi nécessaire pour considérer ces nodules comme véritablement bénins et le bien-fondé d'accorder à ces nodules le même suivi que celui réservé aux nodules à cytologie bénigne. Les principales caractéristiques des études abordées dans cette section sont décrites au [tableau D-2](#) en annexe.

Une étude rétrospective multicentrique a été menée dans le but de déterminer le taux de chirurgies chez un groupe de patients qui présentaient un nodule à cytologie indéterminée GEC bénin (≥ 1 cm) dont les données avaient été recueillies durant une période d'au moins 3 ans [Sipos *et al.*, 2016]. La durée médiane du suivi avait été de 26 mois (étendue : 0 à 44 mois). Au total, 17 % des patients (17/98) ont été soumis à une résection chirurgicale. Dans la majorité des cas (88 %), la chirurgie a été pratiquée durant les 2 premières années suivant le résultat GEC bénin : 10 cas la première année, 5 la deuxième et 2 la troisième. La croissance nodulaire rapide et la taille du nodule étaient les indications les plus souvent citées pour la chirurgie réalisée. Le diagnostic histologique final des spécimens réséqués n'a pas été récupéré pour l'analyse et, par conséquent, le taux de faux négatifs n'a pas été déterminé. Les auteurs ont conclu que la décision de ne pas opérer avait été durable pour la plupart des patients durant la période d'observation de 3 ans de l'étude. La diminution progressive du nombre annuel de chirurgies combinée au taux élevé de chirurgies durant les deux premières années suggère, selon les auteurs, qu'un suivi de plus longue durée est peu susceptible d'affecter significativement le taux de chirurgies pratiquées sur ces nodules. La collecte de données et l'analyse statistique ont été accomplies par le commanditaire, Veracyte Inc.

En guise de comparaison, la proportion de nodules GEC bénins qui ont été réséqués dans l'ensemble des 17 études présentées au tableau 6 a été de 14 % (215/1 495 nodules GEC bénins). La durée médiane du suivi a cependant rarement été précisée. L'étude de Deaver et ses collaborateurs, dans laquelle le taux de résection des nodules GEC bénins a été de 7 %, a rapporté un suivi médian de 46 et 62 mois pour les nodules GEC bénins à cytologie de catégories III et IV, respectivement [Deaver *et al.*, 2018].

Une étude rétrospective financée par Veracyte Inc. a comparé les résultats des patients qui présentaient un nodule (≥ 1 cm) à cytologie indéterminée GEC bénin ($n = 201$) à ceux des patients au diagnostic cytologique bénin ($n = 603$) [Singer *et al.*, 2016].

L'information relative au suivi individuel a été obtenue à partir d'une banque de données de réclamations pour frais médicaux du système de soins de santé américain.

Les patients des 2 groupes ont été appariés dans un ratio de 1:3 sur la base de critères associés au risque relatif de malignité et susceptibles, par conséquent, d'affecter la prise

en charge ultérieure des nodules (genre, âge, année de la biopsie et taille du nodule). La durée minimale permise par l'étude pour le suivi de chaque patient a été de 12 mois après la biopsie (médiane de 19 mois dans chaque groupe). Globalement, la prise en charge des deux groupes de patients a été similaire (tableau 10). Aucune différence significative n'a été observée quant au nombre moyen de visites de suivi ($p = 0,364$) ou d'examen échographiques ($p = 0,811$). Une proportion semblable de patients dans chaque groupe ont eu au moins une visite de suivi ($p = 0,830$) et un examen échographique ($p = 0,707$). Le taux de chirurgies thyroïdiennes ($p = 0,594$) et le temps écoulé entre la biopsie et la chirurgie ($p = 0,721$) n'ont pas été significativement différents. Le diagnostic histologique final des spécimens réséqués et le taux de faux négatifs pour chaque groupe n'ont pas été rapportés. Les auteurs ont mentionné qu'une prise en charge des nodules à cytologie indéterminée GEC bénins comparable à celle des nodules à cytologie bénigne suggère que les cliniciens ont perçu un risque de malignité similaire dans les deux groupes.

Tableau 10 Prise en charge des nodules à cytologie indéterminée GEC bénins en comparaison avec celle des nodules à cytologie bénigne

Étude	Paramètres évalués	Cytologie bénigne	Cytologie indéterminée GEC bénin	p
Singer <i>et al.</i> , 2016 Rétrospective, multicentrique		n = 603	n = 201	
	Suivi moyen (médian), mois*	20,1 (18,8)	20,3 (18,6)	-
	Visites, nombre moyen	5,71	6,09	0,364
	Examen échographique, nombre moyen	0,99	1,01	0,811
	≥ 1 visite de suivi, %	96,4	96,0	0,830
	≥ 1 examen échographique, %	61,7	60,2	0,707
	Chirurgie thyroïdienne, %	10,1	11,4	0,594
	Délai moyen (médian) à la chirurgie, jours	247 (201)	167 (122)	0,721
Angell <i>et al.</i> , 2015 Rétrospective	Hémithyroïdectomie, %	5,3	7,0	0,380
	Thyroïdectomie (totale ou subtotale), %	4,5	4,5	1,000
		n = 1224	n = 58	
	Suivi médian, mois* (étendue)	14,3 (4-40)	13,0 (4-40)	0,21
	Croissance nodulaire significative, %			
	- augmentation (min 2 dimensions) ≥ 20 %	8,3	8,6	0,80
	- augmentation de volume ≥ 50 %	13,8	17,2	0,44
	Répétition de la biopsie, %	16,8	6,9	< 0,05
	Chirurgie thyroïdienne, %	0,9	13,8	< 0,0001
	Malignité, %	33,3	12,5	0,16

GEC : *gene expression classifier*.

* Durée du suivi à partir du moment de la biopsie initiale pour les deux groupes.

Enfin, Angell et ses collaborateurs ont mené une étude rétrospective dans le but d'évaluer la croissance et les caractéristiques échographiques des nodules à cytologie indéterminée durant la période d'observation qui suit un résultat GEC bénin. Les données du suivi échographique étaient disponibles pour 58 des 90 nodules à cytologie indéterminée GEC bénins (64 %) identifiés dans leur centre entre 2010 et 2014. Les données recueillies ont été comparées à celles d'un groupe de nodules à cytologie bénigne (n = 1 224) qui avaient eu une durée de suivi semblable (13,0 vs 14,3 mois; $p = 0,21$). Les caractéristiques de base des patients (âge et genre) et la taille des nodules ont été réparties de façon similaire entre les deux groupes. En revanche, la

proportion de cas de multinodularité a été plus importante parmi les nodules à cytologie bénigne (65,8 vs 50,0 %; $p = 0,02$) et une composition solide a été décrite dans une plus grande proportion de nodules à cytologie indéterminée GEC bénins (60,4 vs 36,7 %; $p < 0,001$). À noter que les nodules à cytologie bénigne ont été évalués au cours d'une période différente (2001-2010) de celle des nodules à cytologie indéterminée GEC bénins. De plus, l'emploi de l'analyse Afirma GEC pour évaluer les nodules à cytologie indéterminée était à la discrétion du clinicien.

La croissance des nodules a été examinée selon deux définitions : augmentation d'au moins 20 % dans un minimum de 2 dimensions et augmentation du volume d'au moins 50 %. Aucune différence significative de la croissance n'a été observée entre les nodules à cytologie indéterminée GEC bénins et les nodules à cytologie bénigne (peu importe la définition employée). Le risque de croissance des nodules à cytologie indéterminée GEC bénins est demeuré similaire à celui des nodules à cytologie bénigne après un ajustement selon la multinodularité, le contenu cystique et la durée du suivi (RC : 1,17 [IC 95 % 0,45-3,10]; $p = 0,75$ selon la première définition et RC : 1,37 [IC 95 % 0,67-2,82]; $p = 0,39$ selon la deuxième). Les nodules à cytologie indéterminée GEC bénins ont ensuite été examinés pour repérer la présence de caractéristiques échographiques suspectes (hypoéchogénicité, dimension plus haute que large, calcification et augmentation du ratio des composantes solides/cystiques). Au total, 5 nodules (8,6 %) ont présenté une de ces caractéristiques (3 cas de dimension plus haute que large et 2 cas hypoéchogènes); 2 des nodules à la dimension plus haute que large ont été réséqués et se sont avérés bénins à l'examen histopathologique. Aucune évaluation similaire portant sur les caractéristiques échographiques des nodules à cytologie bénigne n'a été réalisée.

La prise en charge des deux groupes a aussi été examinée afin de déceler toute différence qui aurait pu survenir au cours de la période de suivi de l'étude. Une reprise de la BAAF a été pratiquée pour 4 nodules du groupe GEC bénin (6,9 %) entre 6 et 37 mois après l'évaluation moléculaire : le résultat a été bénin pour 2 nodules et indéterminé pour les 2 autres. Par ailleurs, une plus grande proportion de nodules ont été soumis à une reprise de la BAAF dans le groupe initialement à cytologie bénigne (16,8 %; $p < 0,05$) : le résultat a été de nouveau bénin pour 190 nodules (92 %), indéterminé pour 8 nodules (4 %), *non-diagnostic* pour 5 nodules (2 %), à suspicion de malignité pour 2 nodules (1 %) et malin pour 1 nodule (< 1 %). Après une durée médiane de 9 mois (6 à 19 mois) suivant le résultat GEC bénin, 13,8 % des nodules à cytologie indéterminée ont été réséqués ($n = 8$); le diagnostic histopathologique a révélé la présence d'un seul cas de malignité (12,5 %). Dans le groupe de nodules à cytologie bénigne, la résection chirurgicale a été moins fréquente (0,9 %; $p < 0,0001$). Parmi les nodules réséqués, une plus grande proportion de nodules malins ont été identifiés (33 %; $p = 0,16$). Il est à noter que le taux de chirurgies observé dans cette étude pour les nodules à cytologie bénigne a été considérablement inférieur à celui rapporté dans la méta-analyse de Bongiovanni, qui incluait 15 104 nodules à cytologie bénigne dont 10,4 % ont été réséqués et 3,7 % avaient une histologie maligne [Bongiovanni *et al.*, 2012].

Les auteurs ont conclu que les nodules à cytologie indéterminée GEC bénins ont une croissance comparable à celle des nodules à cytologie bénigne, ce qui suggère un comportement clinique semblable. Selon eux, une prise en charge similaire est jugée appropriée.

EN BREF

Le taux historique de nodules à cytologie bénigne réséqués a été évalué à 10,4 %; la prévalence du diagnostic histologique malin parmi ces nodules a été de 3,7 % (méta-analyse de Bongiovanni [2012]).

La comparaison entre les nodules à cytologie indéterminée moléculairement bénins et ceux à cytologie bénigne est limitée à deux études réalisées avec la version GEC d'Afirma.

Dans l'une des études, le taux de chirurgies parmi les nodules GEC bénins a été semblable à celui observé pour les nodules à cytologie bénigne (11,4 % vs 10,1 %, respectivement; $p = 0,594$).

Dans l'autre étude, une différence statistiquement significative de nodules réséqués a été rapportée (13,8 % vs 0,9 %, respectivement; $p < 0,0001$); la prévalence de la malignité a été de 12,5 % vs 33,3 %, respectivement ($p = 0,16$).

2.6. Portrait de la prise en charge actuelle des nodules thyroïdiens à cytologie indéterminée

2.6.1. Recommandations internationales

Plusieurs associations ont élaboré des recommandations cliniques relativement à la prise en charge des nodules thyroïdiens à cytologie indéterminée [Meltzer *et al.*, 2019; NCCN, 2019; Gharib *et al.*, 2016; Haugen *et al.*, 2016; Mitchell *et al.*, 2016; Perros *et al.*, 2014]. Un résumé des principales recommandations est présenté au [tableau 11](#). De façon générale, les associations ont souligné l'importance de considérer l'ensemble des caractéristiques cliniques, échographiques et cytologiques pour l'évaluation du risque de malignité et la prise en charge qui en résulte. Une discussion avec le patient concernant les options diagnostiques/thérapeutiques disponibles est habituellement encouragée afin de tenir compte de ses préférences. Les options incluent une reprise de la BAAF (catégorie III), la surveillance active, la lobectomie diagnostique et l'évaluation moléculaire.

Selon le NCCN, les nodules à cytologie indéterminée peuvent faire l'objet d'une surveillance active si l'utilisation d'un test moléculaire donné prédit un risque de malignité comparable à celui observé pour un résultat de cytologie bénin (taux de faux négatifs $\leq 5\%$) [NCCN, 2019]. Certains groupes d'experts ont cependant exprimé leur préoccupation quant à la décision de ne pas réséquer ces nodules [Paschke *et al.*, 2017; Ferris *et al.*, 2015; Bernet *et al.*, 2014]. Les données sur le suivi à long terme des nodules à cytologie indéterminée qui évitent la chirurgie diagnostique en raison d'un résultat moléculairement bénin ont été jugées insuffisantes. De plus, étant donné que la valeur prédictive des tests diagnostiques est tributaire de la prévalence de la maladie dans chaque catégorie cytologique et que celle-ci varie d'un établissement à l'autre, la performance des tests moléculaires est susceptible de varier considérablement.

L'American Thyroid Association indique que, lorsque l'évaluation moléculaire est considérée, les patients doivent être informés des avantages et limites potentiels du test réalisé ainsi que des incertitudes associées aux implications cliniques du résultat obtenu [Haugen *et al.*, 2016].

Tableau 11 Recommandations en lien avec la prise en charge des nodules thyroïdiens à cytologie indéterminée

Source	Catégorie diagnostique de la BAAF (Bethesda)	
	Atypie de signification indéterminée OU lésion folliculaire de signification indéterminée (III)	Néoplasme folliculaire OU suspicion d'un néoplasme folliculaire (IV)
AHNS [Meltzer <i>et al.</i> , 2019]	Reprise de la BAAF, surveillance ou chirurgie. Si la reprise indique à nouveau une cytologie de catégorie III, considérer l'évaluation moléculaire, la surveillance ou la chirurgie.	Considérer l'évaluation moléculaire, la surveillance ou la chirurgie.
NCCN [2019]	En présence de caractéristiques cliniques/échographiques à suspicion élevée de malignité, les options incluent la lobectomie ou la thyroïdectomie totale* pour un diagnostic/traitement définitif. En l'absence de caractéristiques cliniques/échographiques à suspicion élevée de malignité, les options incluent : la lobectomie diagnostique, le diagnostic moléculaire, une reprise de la BAAF ou la surveillance active. Si le Dx moléculaire prédit un risque de malignité comparable à celui observé à la suite d'une cytologie bénigne[†], la surveillance active peut être considérée. Si le Dx moléculaire suggère la présence d'un nodule malin, les options incluent la lobectomie ou la thyroïdectomie totale pour un diagnostic/traitement définitif.	En l'absence de caractéristiques cliniques/échographiques à suspicion élevée de malignité, les options incluent : la lobectomie diagnostique, le diagnostic moléculaire ou la surveillance active. Si le Dx moléculaire prédit un risque de malignité comparable à celui observé à la suite d'une cytologie bénigne[†], la surveillance active peut être considérée. Dans certains cas (selon les facteurs de risque cliniques/échographiques présents et la préférence du patient), une lobectomie peut être pratiquée pour un diagnostic/traitement définitif. Si le Dx moléculaire suggère la présence d'un nodule malin, les options incluent la lobectomie ou la thyroïdectomie totale pour un diagnostic/traitement définitif.
	(Catégorie 2A : Faible niveau de preuve; le consensus est uniforme concernant le caractère approprié de l'intervention)	
ATA [Haugen <i>et al.</i> , 2016]	Après avoir considéré les caractéristiques cliniques et échographiques suspectes, une reprise de la BAAF ou une analyse moléculaire peuvent être utilisées pour compléter l'évaluation du risque de malignité plutôt que de procéder directement avec une stratégie de surveillance ou de chirurgie diagnostique. La préférence du patient et la faisabilité devraient être considérées dans la prise de décisions cliniques. (Faible recommandation; preuve de qualité modérée) Si la reprise de la BAAF, l'analyse moléculaire ou les deux ne sont pas réalisées ou n'apportent pas de résultat concluant, la surveillance ou la chirurgie diagnostique peuvent être pratiquées, en fonction des facteurs de risque cliniques, des caractéristiques échographiques et de la préférence du patient. (Forte recommandation; preuve de faible qualité)	La chirurgie diagnostique représente la norme établie de longue date. Cependant, après avoir considéré les caractéristiques cliniques/échographiques, une analyse moléculaire peut être utilisée pour compléter l'évaluation du risque de malignité plutôt que de procéder directement à la chirurgie. La préférence du patient et la faisabilité devraient être considérées dans la prise de décisions cliniques. (Faible recommandation; preuve de qualité modérée) Si l'analyse moléculaire n'est pas réalisée ou n'apporte pas de résultat concluant, la chirurgie peut être considérée pour le retrait et le diagnostic définitif du nodule. (Forte recommandation; preuve de faible qualité)

Catégorie diagnostique de la BAAF (Bethesda)		
Source	Atypie de signification indéterminée OU lésion folliculaire de signification indéterminée (III)	Néoplasme folliculaire OU suspicion d'un néoplasme folliculaire (IV)
AACE/ACE/AME [Gharib <i>et al.</i> , 2016]	Considérer une prise en charge conservatrice dans le cas de caractéristiques cliniques favorables (antécédents personnels et familiaux, taille de la lésion et caractéristiques échographiques ou d'élastographie de faible risque). (Niveau de preuve 3; recommandation grade B)[‡] Répéter la BAAF et solliciter l'avis d'un cytopathologiste expérimenté. (Niveau de preuve 3; recommandation grade B)[‡] La biopsie au trocart peut être considérée, mais son utilisation courante n'est pas recommandée. (Niveau de preuve 4; recommandation grade C)[‡] Aucune recommandation en faveur ou contre l'évaluation moléculaire de routine. (Niveau de preuve 3; recommandation grade D)[‡]	La chirurgie est recommandée pour la plupart des lésions de cette catégorie. (Niveau de preuve 2; recommandation grade A)[‡] La lobectomie avec isthmectomie est recommandée; la thyroïdectomie totale peut être pratiquée, selon la situation clinique, la coexistence d'un nodule bilatéral ou la préférence du patient. (Niveau de preuve 2; recommandation grade A)[‡] Les coupes congelées ne sont habituellement pas utiles dans ce groupe. (Niveau de preuve 3; recommandation grade D)[‡] Considérer un suivi clinique sous étroite surveillance chez une minorité de cas qui présentent des caractéristiques cliniques et échographiques favorables, après une consultation multidisciplinaire et une discussion avec chaque patient sur les options de traitement. (Niveau de preuve 4; recommandation grade C)[‡]
Lignes directrices multidisciplinaires nationales (Royaume-Uni) [Mitchell <i>et al.</i> , 2016] Basé sur les recommandations de la BTA [Perros <i>et al.</i> , 2014]	Répéter l'échographie et la BAAF. Si la reprise indique à nouveau une cytologie de catégorie III, consulter l'équipe multidisciplinaire et considérer l'hémi-thyroïdectomie (résection du lobe thyroïdien et de l'isthme) diagnostique.	Hémi-thyroïdectomie (résection du lobe thyroïdien et de l'isthme) diagnostique. Considérer la thyroïdectomie totale pour les lésions > 4 cm, qui présentent un risque de malignité plus élevé.

AACE : American Association of Clinical Endocrinologists; ACE : American College of Endocrinology; AHNS : American Head and Neck Society; AME : Associazione Medici Endocrinologi; ATA : American Thyroid Association; BTA : British Thyroid Association; Dx : diagnostic; ETA : European Thyroid Association; NCCN : National Comprehensive Cancer Network.

* La thyroïdectomie totale peut être considérée pour les néoplasmes à cellules de Hürthle, en présence d'antécédents d'exposition au rayonnement ionisant ou d'une lésion controlatérale.

† Implique un taux de faux négatifs comparable à celui observé pour un résultat de cytologie de catégorie II, soit ≤ 5 %.

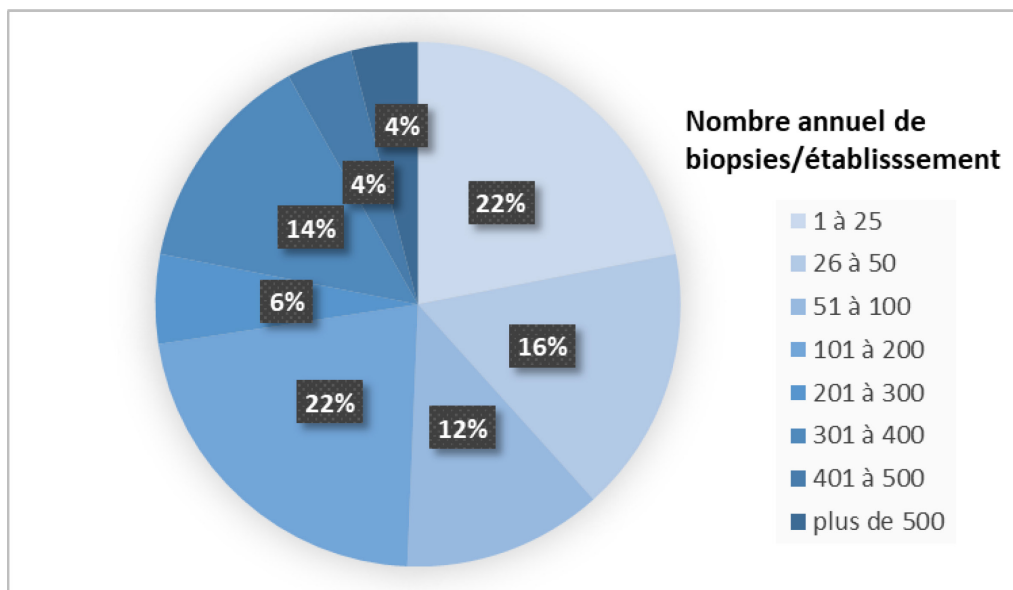
‡ Se référer à la table 1 de l'article publié par Gharib et ses collaborateurs [2016] pour l'interprétation du niveau de preuve et du grade de la recommandation.

2.6.2. Contexte québécois

Un algorithme illustrant la prise en charge actuelle des nodules thyroïdiens selon le résultat de la cytologie est détaillé en annexe ([figure B-1](#)). En l'absence de facteurs de risque cliniques et échographiques élevés, les nodules à cytologie de catégorie Bethesda III sont généralement soumis à une reprise de la BAAF; la lobectomie diagnostique dans cette catégorie est habituellement pratiquée à la suite d'un second résultat de cytologie indéterminée. Pour les nodules de la catégorie Bethesda IV, la lobectomie représente la pratique standard. Advenant la décision d'introduire les tests moléculaires dans la pratique québécoise, le suivi réservé aux nodules à cytologie indéterminée moléculairement bénins serait comparable à celui des nodules à cytologie bénigne qui ne présentent pas de risque élevé. Au Québec, ces nodules bénéficient d'un suivi échographique tous les 12 à 24 mois; l'absence de croissance significative définie comme une augmentation d'au moins 20 % dans un minimum de 2 dimensions ou une augmentation du volume d'au moins 50 % permet de mettre fin au suivi après une période de 4 à 5 ans.

Dans le but de déterminer le volume des BAAF et des chirurgies thyroïdiennes pratiquées au Québec annuellement, la base de données SMOD de la RAMQ a été interrogée au moyen des codes de rémunération à l'acte employés par les médecins spécialistes. Entre 2009 et 2018, le nombre annuel de BAAF de la thyroïde est passé de 8 622 à 13 880. La compilation du nombre d'actes enregistrés par établissement en 2018 a montré que la moitié des établissements ont fait moins de 100 BAAF au cours de cette période; en revanche, une minorité d'établissements (8 %) en ont réalisé plus de 400 durant la même année. Ces observations suggèrent la présence d'écarts importants quant au volume d'examen cytopathologiques effectués dans chaque établissement.

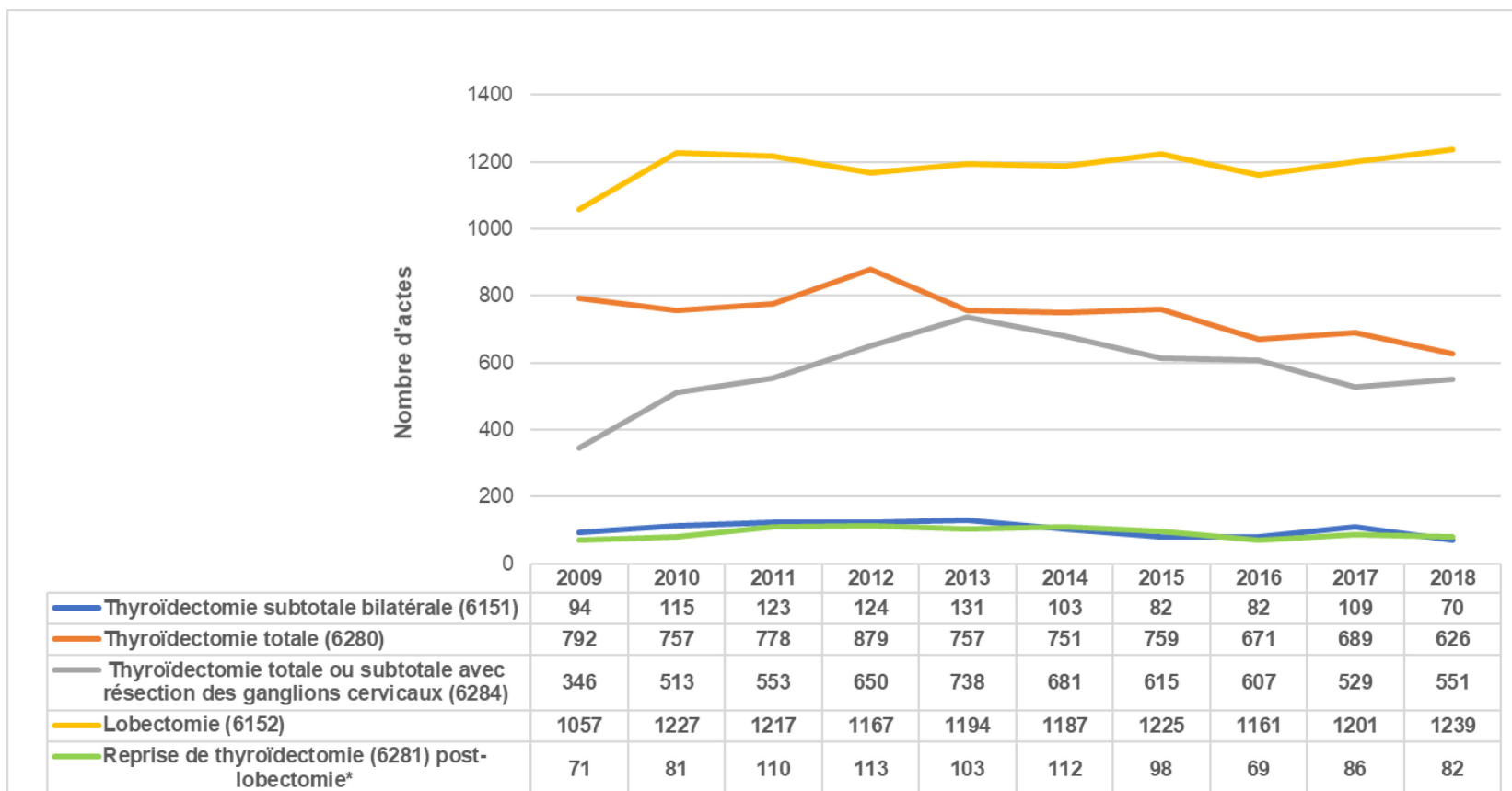
Figure 2 Volume de biopsies par aspiration à l'aiguille fine pratiquées par établissement en 2018 au Québec*



* Données disponibles pour 73 établissements et représentant 84 % des actes de biopsie (0237) pratiqués en 2018.

Le nombre d'actes enregistrés entre 2009 et 2018 selon le type de chirurgie thyroïdienne réalisée est représenté à la figure 3. Les données révèlent qu'au cours de la période évaluée la lobectomie a été l'intervention chirurgicale la plus souvent pratiquée en comparaison avec la thyroïdectomie totale ou subtotale et la reprise de la thyroïdectomie. En 2018, 1 239 lobectomies ont été réalisées; il n'est cependant pas possible de déterminer quelle proportion de ces actes ont été accomplis dans un objectif diagnostique sur des nodules à cytologie indéterminée. Au cours des dix dernières années, le nombre annuel de lobectomies est demeuré relativement stable.

Figure 3 Nombre d'actes enregistrés au Québec entre 2009 et 2018 selon le type de chirurgie thyroïdienne (sans égard au résultat de la cytologie)



* Sur les 925 actes de reprise (postlobectomie) enregistrés entre 2009 et 2018, 143 ont été pratiqués à la suite d'une lobectomie réalisée avant 2009. De plus, plusieurs actes de reprise enregistrés entre 2009 et 2018 n'avaient pas été précédés d'un acte de chirurgie thyroïdienne (ni avant, ni après 2009); ces actes de reprise (n = 346) n'ont pas été considérés dans l'analyse des données.

2.7. Perspective des patients

Deux études qui avaient évalué la perspective des patients dans le contexte spécifique à la prise en charge des nodules thyroïdiens à cytologie indéterminée ont été repérées [Wong *et al.*, 2020; Lee *et al.*, 2017]. Dans celle menée par Lee et ses collaborateurs, des scénarios hypothétiques ont été présentés à deux groupes de patients suivis dans le réseau de santé universitaire de la région de Toronto : ceux présentant un nodule thyroïdien à cytologie indéterminée (groupe A; n = 50) et ceux orientés vers la clinique d'otorhinolaryngologie pour d'autres motifs (groupe B; n = 50). La préférence des patients pour la chirurgie, l'observation et l'utilisation de tests moléculaires a été déterminée en fonction du risque de malignité, du coût et des propriétés diagnostiques des tests moléculaires (tableau 12).

Les patients devaient initialement indiquer leur préférence entre deux options :

1) l'observation avec reprise de l'échographie et de la biopsie ou 2) la lobectomie diagnostique avec possibilité de reprise pour compléter la thyroïdectomie. Le risque de malignité a d'abord été fixé à 50 % et il a varié ensuite à la hausse et à la baisse jusqu'à ce que le patient modifie sa préférence. Dans les scénarios comprenant les tests moléculaires (Afirma GEC, ThyroSeq v0 ou les deux), la proportion de patients qui ont préféré ces derniers à une prise en charge standard (observation ou chirurgie) a été évaluée. Chaque scénario incluait une variation des coûts de l'analyse pour le patient (aucun coût vs coût véritable de l'analyse) ainsi qu'une variation des propriétés diagnostiques des tests offerts (hypothétiquement parfaites vs propriétés véritables de l'analyse). Les tests hypothétiquement parfaits étaient associés à une sensibilité et une spécificité de 100 %.

Tableau 12 Préférence des patients pour la prise en charge des nodules à cytologie indéterminée selon le risque de malignité, le coût et les propriétés diagnostiques des tests moléculaires

Paramètres évalués	Total (n = 100)	Groupe A (n = 50)	Groupe B (n = 50)	p (A vs B)
Seuil du risque de malignité pour préférer la chirurgie, %	38,6 (IC 95 % 33,4-43,7)	33,9	43,2	0,072
% pts préférant l'évaluation moléculaire :				
- sans frais/propriétés réelles	85,0	86,7	83,3	0,419
- sans frais/propriétés parfaites	89,3	-	-	-
- frais applicables*/propriétés réelles	51,7	51,3	52,0	0,908
- frais applicables*/propriétés parfaites	49,5	-	-	-

pts : patients.

* Coûts des tests moléculaires payés par le patient.

Les résultats obtenus ont montré que le seuil moyen au-delà duquel les patients ont préféré la chirurgie à l'observation correspond à un risque de malignité de 38,6 %. Lorsqu'un test moléculaire était proposé gratuitement, 85 % des personnes ont préféré cette option à celle de l'observation ou de la chirurgie. La proportion de patients qui ont choisi l'évaluation moléculaire a cependant diminué significativement pour atteindre

51,7 % lorsque les coûts de l'analyse ont été facturés au patient ($p < 0,001$). Les propriétés diagnostiques des tests proposés ont eu peu d'impact sur la décision des répondants; les tests n'étaient pas préférés davantage en présence de caractéristiques hypothétiquement parfaites, en comparaison avec les propriétés diagnostiques réelles/imparfaites (89,3 vs 85,0 %; $p = 0,11$ lorsque sans frais pour le patient et 49,5 vs 51,7 %; $p = 0,54$ lorsque les coûts étaient facturés au patient). Les patients ont préféré les tests moléculaires lorsqu'ils étaient sans frais pour eux, peu importe leurs propriétés diagnostiques (85,0 vs 51,7 %; $p < 0,001$ selon les propriétés réelles et 89,3 vs 49,5 %; $p < 0,001$ selon les propriétés hypothétiquement parfaites). Enfin, l'analyse multivariée a révélé que le genre, un diagnostic de nodule à cytologie indéterminée et le niveau de revenu étaient des facteurs de prédiction indépendants de la préférence du patient pour l'évaluation moléculaire.

En résumé, les patients ont préféré l'évaluation moléculaire à la prise en charge habituelle des nodules à cytologie indéterminée. Cette préférence était tributaire du coût des tests offerts, mais non de leurs propriétés diagnostiques. Selon les auteurs, les patients considéreraient les tests moléculaires comme supérieurs à la biopsie pour le diagnostic des nodules thyroïdiens, sans compréhension réelle des concepts associés à leur performance. Les auteurs ont ajouté que l'utilisation judicieuse de ces tests nécessite une décision partagée entre le patient et le clinicien afin de prévenir les décisions résultant d'un mauvais jugement.

L'étude réalisée par Wong et ses collaborateurs avait pour objectif de caractériser l'impact des tests moléculaires sur la qualité de vie des patients qui présentent un nodule à cytologie indéterminée [Wong *et al.*, 2020]. Les patients admissibles étaient ceux dont le nodule avait été soumis à un examen cytologique (BAAF) entre mai 2016 et juin 2017. Suivant un résultat indéterminé de la cytologie, une évaluation moléculaire avec Afirma GEC ou ThyroSeq v2.1 a été réalisée (emploi systématique); les données relatives à la prise en charge subséquente et au diagnostic histologique des spécimens réséqués ont été publiées séparément et rapportées à la section 2.4 de cet avis [Livhits *et al.*, 2018]. Une version abrégée du questionnaire ThyPro-39 a été envoyée à tous les patients préalablement informés du résultat de l'évaluation cytologique et moléculaire, le cas échéant. Quatre catégories de question ont été analysées : symptômes du goitre, anxiété, dépression et altération de la vie quotidienne. L'expérience associée à chacune de ces catégories a été qualifiée par le patient à l'aide des options suivantes (incluant le nombre de points alloués pour chaque réponse) : pas du tout (0), minimalement (1), un peu (2), assez (3) et beaucoup (4). Un score a été calculé pour chaque catégorie sur une échelle de 0 à 100, de sorte que les scores les plus élevés représentaient une plus grande détérioration de la qualité de vie. Les données des patients soumis à l'évaluation moléculaire ont été analysées sans égard au type de test réalisé (Afirma GEC ou ThyroSeq v2.1). Les principaux résultats de l'étude sont présentés ci-dessous.

Tableau 13 Impact de l'évaluation moléculaire sur la qualité de vie des patients

ThyPro-39	Bethesda II (n = 251)	Bethesda V/VI (n = 23)	Bethesda III/IV moléculairement bénin (n = 37)	Bethesda III/IV moléculairement suspect (n = 21)
Symptômes du goitre	13,3 (17,2)	20,6 (20,3)	10,4 (12,9)	20,5 (20,8)*
Anxiété	22,5 (23,6)	34,2 (30,5)†	22,8 (19,2)	25,7 (27,0)
Dépression	27,7 (20,7)	32,1 (17,8)	21,0 (14,9)	33,3 (24,7)*
Altération de la vie quotidienne	12,0 (19,6)	25,8 (28,4)†	11,6 (16,6)	8,3 (16,6)

Données présentées sous forme de moyenne (écart type).

* Différence statistiquement significative entre les groupes moléculairement suspect et bénin.

† Différence statistiquement significative entre les groupes de cytologie maligne (ou à suspicion de malignité) et bénigne.

Parmi les patients qui ont consenti à répondre au questionnaire, les résultats de cytologie étaient répartis de la façon suivante : 76 % Bethesda II (n = 251), 7 % Bethesda V/VI (n = 23) et 17 % Bethesda III/IV (n = 58, incluant 37 nodules moléculairement bénins et 21 nodules moléculairement suspects). Aucune différence significative n'a été observée pour chacune des quatre catégories de question évaluées entre les patients porteurs d'un nodule à cytologie bénigne et ceux porteurs d'un nodule à cytologie indéterminée moléculairement bénin. Par ailleurs, les patients qui présentaient un nodule à cytologie bénigne ont montré une meilleure qualité de vie en comparaison avec ceux dont la cytologie était maligne (ou à suspicion de malignité) en termes d'anxiété (p = 0,026) et d'altération de la vie au quotidien (p = 0,001). Toutefois, aucune différence significative n'a été observée entre ces deux groupes quant aux symptômes du goitre et de la dépression. Parmi les patients présentant un nodule à cytologie indéterminée, ceux au résultat moléculairement bénin ont rapporté moins de symptômes du goitre (p = 0,033) et de la dépression (p = 0,026) que ceux au résultat moléculairement suspect, alors qu'aucune différence n'a été observée entre les deux groupes en lien avec l'anxiété et l'altération de la vie quotidienne. Étant donné que l'emploi du test était systématique pour tous les nodules à cytologie indéterminée, l'impact des tests moléculaires dans la catégorie de cytologie indéterminée (avec et sans) n'a pu être évalué.

Bien que le nombre de nodules à cytologie indéterminée ait été limité dans cette étude, les auteurs ont suggéré qu'un résultat bénin au test moléculaire serait aussi rassurant pour les patients qu'un résultat de cytologie bénigne. Un suivi prolongé est cependant nécessaire pour déterminer l'impact à long terme des tests moléculaires sur la qualité de vie des patients avec un suivi échographique régulier.

2.8. Consultation - Perspective des experts

Les données scientifiques, expérientielles et contextuelles ont fait l'objet d'échanges avec les membres du comité consultatif et du Comité de l'évolution des pratiques en oncologie. Les membres sont d'avis que le besoin de santé lié au diagnostic des nodules thyroïdiens à cytologie indéterminée est important et que le remboursement des analyses moléculaires évaluées aurait vraisemblablement des retombées cliniques et organisationnelles favorables.

De l'avis de certains experts consultés, l'effet de la prévalence de la maladie sur la valeur prédictive négative du test et le risque résiduel d'un nodule moléculairement bénin ne constitue pas un élément préoccupant. Le suivi clinique/échographique réservé aux nodules moléculairement bénins, couplé à la progression lente des lésions malignes de stade précoce, permettent d'anticiper que les nodules faussement négatifs qui auront échappé à la détection moléculaire seront décelés, le cas échéant, sans que cela porte un important préjudice aux patients concernés.

Par ailleurs, l'impact de la prévalence de la maladie peut être atténué par l'application de critères d'admissibilité prédéfinis qui réduisent le niveau de risque. Selon les experts consultés, le test ne devrait pas être utilisé de façon systématique sur tous les nodules à cytologie indéterminée; l'emploi devrait être fait uniquement sur une base sélective. La population ciblée devrait être limitée aux patients qui présentent un nodule à cytologie de catégorie Bethesda III (après une 2^e BAAF) ou IV, dont la taille est comprise entre 1 et 4 cm et qui ne présentent pas de facteurs de risque cliniques ou échographiques élevé. Les patients qui répondent d'emblée à une indication pour la chirurgie, de même que ceux qui manifestent une préférence pour la chirurgie ou l'observation, ne devraient pas se voir offrir une évaluation moléculaire. Enfin, l'utilisation devrait être restreinte aux patients chez qui le résultat du test est susceptible d'influer sur la prise en charge.

2.9. Volet économique

2.9.1. Littérature scientifique

Plusieurs études traitant de l'aspect économique associé à l'utilisation des tests moléculaires pour préciser le diagnostic des nodules à cytologie indéterminée ont été répertoriées. La sélection a cependant été restreinte à celles dont le scénario de base était représentatif de la population ciblée et de la prise en charge actuelle des nodules thyroïdiens. Les études qui ont inclus (sans distinction) les nodules qui présentent des facteurs de risque cliniques ou échographiques élevé n'ont pas été retenues, parce que ces nodules ne seraient pas soumis à une évaluation moléculaire. Par ailleurs, comme la lobectomie est le seul type de chirurgie recommandé pour les nodules à cytologie indéterminée en l'absence de facteurs de risque élevé ([annexe B](#)), que la valeur prédictive positive des tests moléculaires est limitée et que les données présentement disponibles ne permettent pas d'employer le profil moléculaire pour déterminer l'étendue de la chirurgie, les études dans lesquelles un résultat moléculairement suspect conduit à une thyroïdectomie totale n'ont pas été sélectionnées. De même, les études dont le

modèle comprend une reprise de thyroïdectomie systématique pour tous les cas de cancer décelés après lobectomie (qu'elle soit diagnostique ou la conséquence d'un résultat moléculairement suspect) ont été exclues parce que les lignes directrices actuelles ne le recommandent la reprise de thyroïdectomie qu'en présence de certains éléments⁶. Les résultats de l'évaluation de la qualité des études économiques sélectionnées sont présentés en annexe ([tableau D-3](#)).

Zanocco et ses collaborateurs ont évalué l'efficacité des tests moléculaires en comparaison avec la lobectomie diagnostique chez des patients qui présentaient un nodule à cytologie indéterminée dont le risque avait été stratifié en fonction des caractéristiques échographiques [Zanocco *et al.*, 2020]. Une base de données institutionnelle a été interrogée pour repérer les nodules soumis à l'évaluation moléculaire avec l'analyse Afirma GEC entre 2012 et 2016. Ces nodules ont ensuite été assignés à une catégorie de risque échographique selon la classification adoptée par l'ATA; le taux de résultats moléculairement bénins a été calculé pour chaque catégorie. Selon le modèle décisionnel généré, les nodules moléculairement suspects étaient réséqués par lobectomie et les nodules au profil bénin étaient placés en observation annuelle. Un horizon temporel de 40 ans a été fixé afin d'approximer l'espérance de vie résiduelle d'un patient de 40 ans. L'efficacité a été mesurée en QALY (*quality-adjusted life year*) et la stratégie optimale a été définie comme celle produisant la meilleure efficacité incrémentale avec un rapport coût-efficacité incrémental inférieur à 100 000 \$/QALY.

Tableau 14 Évaluation moléculaire et rapport coût-efficacité en fonction du risque échographique

Suspicion échographique	Nodules à cytologie indéterminée, n (%)	Taux de résultats GEC bénin (%)	Coût* incrémental de l'analyse GEC (\$)	Efficacité incrémentale de l'analyse GEC (QALY)	Rapport coût-efficacité incrémental de l'analyse GEC (\$/QALY)
Élevée	12 (9,8)	8,3	3 300	0,0068	490 000
Intermédiaire	49 (40)	35	2 400	0,029	84 000
Faible / Très faible	62 (50)	44	2 100	0,036	58 255

GEC : *gene expression classifier*; QALY : *quality-adjusted life year*.

* Le coût en dollars américains de la lobectomie (5 252 \$), de ses complications (lésion NLR : 13 132 \$) et de l'hypothyroïdie (193 \$), de l'analyse moléculaire (3 600 \$) et du suivi échographique (131 \$) a été considéré lorsqu'il était jugé applicable.

⁶ Une reprise de thyroïdectomie pour goitre restant (chirurgie de complétion) est pratiquée lorsqu'un carcinome papillaire est décelé suivant la lobectomie et qu'un des éléments suivants est identifié : la tumeur fait plus de 4 cm, les marges de résection sont positives, en présence d'extension extrathyroïdienne macroscopique, d'une atteinte ganglionnaire, d'une maladie controlatérale, d'une invasion vasculaire ou d'un cancer peu différencié [NCCN, 2019].

Au total, 123 nodules ont été répertoriés. Une minorité (9,8 %) présentaient une suspicion échographique élevée, 40 % étaient classés dans la catégorie intermédiaire et 50 % dans la catégorie à faible ou très faible risque. L'évaluation moléculaire s'est avérée efficiente lorsque le taux de résultats moléculairement bénins parmi les catégories de suspicion échographique était supérieur à 31 %, tel qu'observé pour les catégories « intermédiaire » et « faible/très faible » (tableau 14). Au seuil de 100 000 \$/QALY, l'évaluation moléculaire est devenue efficiente pour les nodules à suspicion élevée lorsque le taux de complications suivant la lobectomie excédait 14 %, que le coût de l'analyse était inférieur à 963 \$ ou que le facteur d'ajustement de la qualité de vie pour l'hypothyroïdisme était inférieur à 0,93. L'analyse de sensibilité bidirectionnelle de la probabilité d'un résultat moléculairement bénin et du coût de l'observation à la suite d'un résultat moléculairement bénin a révélé que le modèle était peu sensible aux variations du coût de l'observation des nodules à suspicion échographique élevée GEC bénins, alors que l'efficacité de l'évaluation moléculaire dans les autres nodules était très sensible à ces variations. L'analyse de sensibilité probabiliste a révélé que l'utilisation d'Afirma GEC était efficiente dans 78,2 % des cas de faible/très faible suspicion, 55,1 % des cas de suspicion intermédiaire et dans aucun cas de suspicion échographique élevée.

Les auteurs ont conclu que l'évaluation moléculaire des nodules à cytologie indéterminée avec la version GEC d'Afirma était efficiente uniquement en l'absence de caractéristiques échographiques à risque élevé de malignité, parce que le taux de résultats GEC bénins dans cette catégorie était trop faible (8,3 %). Pour les nodules à suspicion échographique intermédiaire et ceux à faible ou très faible suspicion, le BCR (35 et 44 %, respectivement) a été similaire à celui observé dans l'étude de validation clinique originale (41 %) et les rapports coût/efficacité incrémentaux étaient en deçà du seuil d'efficacité. Les auteurs ont présumé que la meilleure performance de la version GSC par rapport à GEC (spécificité : 68 % vs 52 %) devrait se traduire par une amélioration de l'efficacité globale de l'analyse.

Une autre étude a comparé le rapport coût-efficacité associé à une approche diagnostique basée sur la lobectomie ou l'évaluation moléculaire (avec Afirma GSC ou ThyroSeq v3) [Nicholson *et al.*, 2019]. Dans le modèle évalué, les nodules moléculairement suspects ainsi que ceux qui n'avaient produit aucun résultat interprétable étaient réséqués par lobectomie alors que les nodules bénins étaient observés annuellement durant un horizon temporel de 20 ans. Étant donné que la chirurgie thyroïdienne est plutôt sécuritaire et que les cancers différenciés de la thyroïde sont généralement indolents, Nicholson et ses collaborateurs ont privilégié une évaluation de l'efficacité qui ne repose pas sur le gain de QALY. Comme l'objectif de l'analyse moléculaire consiste à préciser le diagnostic afin d'orienter le patient vers une prise en charge optimale, l'efficacité de chaque approche a été représentée par la probabilité de prise en charge optimale basée sur le bon diagnostic. Le coût et l'efficacité du scénario de base ont été résumés sous la forme du coût par bon diagnostic.

Tableau 15 Rapport coût-efficacité des différentes approches diagnostiques

Approche	Coût* / cas (\$)	Probabilité de prise en charge optimale (%)	Coût / bon diagnostic (\$)
Lobectomie	9 602	25,0 [†]	38 408
ThyroSeq v3	10 451	73,2	14 277
Afirma GSC	11 385	63,7	17 873

* Le coût en dollars américains de la lobectomie (9 520 \$), de ses complications (hématome : 6 336 \$; paralysie transitoire ou permanente du nerf laryngé récurrent (NLR) : 2 365 ou 7 044 \$) et de l'hypothyroïdie (1 852 \$), de la reprise d'une biopsie pour l'analyse moléculaire (142 \$), des tests ThyroSeq v3 et Afirma GSC (3 600 \$), d'une visite de suivi clinique (106 \$) et d'un examen échographique (120 \$) a été considéré lorsqu'il était applicable dans les scénarios respectifs.

† Dans le scénario de base, la prévalence de la maladie (nodules malins/NIFTP) parmi les nodules à cytologie indéterminée a été fixée à 25 %.

Les résultats présentés dans le tableau 15 montrent que, bien que la lobectomie diagnostique ait présenté le plus bas coût par cas, l'utilisation des tests moléculaires a été associée au plus bas coût par bon diagnostic. Les auteurs ont expliqué que l'économie réalisée avec les deux tests résulte de la réduction du nombre de chirurgies inutiles pratiquées pour les cas véritablement bénins. L'analyse de sensibilité unidirectionnelle a révélé que ThyroSeq v3 demeure l'option la moins coûteuse par bon diagnostic sur un large éventail de variations du coût de la lobectomie (6 235 \$ à 12 471 \$), de l'analyse ThyroSeq v3 (2 880 \$ à 4 867 \$) et Afirma GSC (2 560 \$ à 7 680 \$), de la sensibilité (TSv3 : 86 à 98 %; GSC : 79 à 98 %) et de la spécificité (TSv3 : 75 à 87 %; GSC : 60 à 76 %), de la prévalence de la maladie (5 à 50 %) et de la durée de l'observation (5 à 20 ans) après un résultat moléculairement bénin. La variation simultanée des coûts des deux tests moléculaires a montré que ThyroSeq v3 demeurerait l'option privilégiée dans les intervalles de coûts testés; la lobectomie n'a jamais été reconnue comme l'option privilégiée en comparaison avec l'évaluation moléculaire. L'analyse de sensibilité probabiliste a indiqué que ThyroSeq v3 était favorisé dans 68,5 % des cas (vs 25,0 % pour Afirma GSC et 6,5 % pour la lobectomie).

Si le coût par bon diagnostic apparaît comme un complément adéquat au coût par QALY pour mesurer l'efficacité dans le présent contexte, le seuil de rentabilité lié à cette stratégie n'est pas connu. Aussi, puisque seuls les coûts propres à chaque approche diagnostique ont été considérés, les données ne devraient pas être interprétées en tant que coût absolu par diagnostic mais plutôt en termes de coût relatif par approche diagnostique.

Balentine et ses collaborateurs ont comparé l'efficacité, en coût par QALY, de l'analyse Afirma GEC à celle de la lobectomie diagnostique pour l'évaluation des nodules à cytologie indéterminée en l'absence de facteurs de risque élevé de malignité [Balentine *et al.*, 2018]. Le scénario de base reposait initialement sur les recommandations de prise en charge publiées par l'ATA en 2009, selon lesquelles un cancer (> 1 cm) détecté après une lobectomie était systématiquement soumis à une thyroïdectomie de complétion. Parallèlement, la majorité des nodules au résultat GEC suspect ont été dirigés vers une thyroïdectomie totale (67 %) alors que les autres (33 %) étaient orientés vers la

lobectomie. Les nodules GEC bénins ont été placés en surveillance échographique. Sur un horizon temporel de 5 ans, le coût et l'efficacité incrémentale d'Afirma GEC a été de 3 300 \$ US et de - 0,03 QALY (rapport coût/efficacité incrémental non défini), ce qui indique que la lobectomie est l'option la moins coûteuse et la plus efficace. Cependant, si le facteur d'ajustement de la qualité de vie selon l'état de santé post-lobectomie (sans hormonothérapie de substitution à long terme) est inférieur à 0,954, il est alors attendu que l'emploi du test moléculaire sera efficient au seuil de 100 000 \$ US/QALY. L'analyse de sensibilité bidirectionnelle a montré que la modélisation de l'efficacité était très sensible aux valeurs d'utilité des états de santé associés à chaque stratégie diagnostique (différence d'utilité) alors que la variation du coût de la surveillance pour différentes valeurs d'utilité a eu peu d'effets.

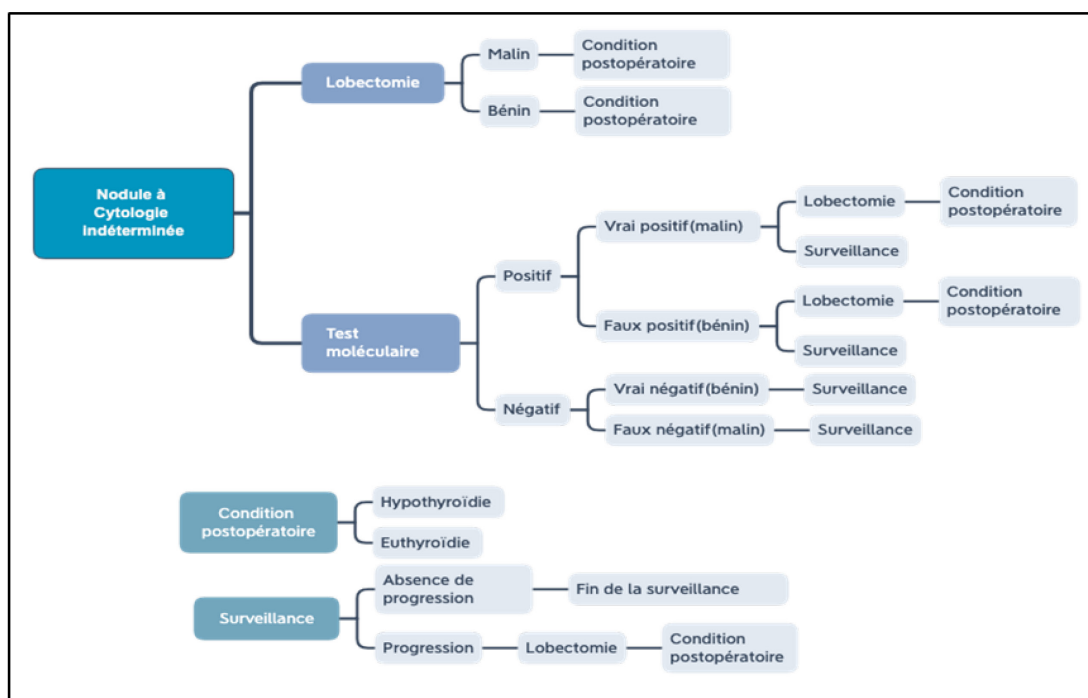
Dans le but de déterminer si l'application des plus récentes recommandations (2015) pouvait influencer sur les conclusions tirées avec le scénario de base, l'analyse a été reprise en considérant la lobectomie comme une approche chirurgicale adéquate en présence d'un cancer détecté post-lobectomie, de même qu'à la suite d'un résultat GEC suspect. Les données obtenues ont montré que le test moléculaire était légèrement plus efficace que la lobectomie (+ 0,0017 QALY), mais beaucoup plus coûteux (+ 3 130 \$ US); conséquemment, le rapport coût/efficacité incrémental (1 843 000 \$ US/QALY) a été considérablement plus élevé que le seuil d'efficacité prédéfini (100 000 \$ US/QALY).

Les auteurs ont conclu que leur analyse prédisait que la lobectomie dominerait l'évaluation moléculaire avec Afirma GEC comme stratégie pour exclure la malignité parmi les nodules à cytologie indéterminée. Ils ont souligné la sensibilité du modèle aux valeurs d'utilité des divers états de santé associés à chaque stratégie. Pour la lobectomie, les valeurs employées sont celles issues de la littérature et générées à partir d'échantillons limités. L'évaluation de l'utilité des états de santé liés à l'observation des nodules à cytologie indéterminée moléculairement bénins n'a fait l'objet d'aucune étude, ce qui a amené les auteurs à s'appuyer sur les données des patients sous surveillance pour un nodule à cytologie maligne ou bénigne. Selon eux, une évaluation rigoureuse des valeurs et préférences des patients dans le contexte de ces états de santé serait souhaitable et pourrait conduire à des conclusions différentes concernant l'efficacité des approches diagnostiques évaluées.

2.9.2. Évaluation économique primaire

Puisque les résultats issus des évaluations de l'efficacité réalisées dans d'autres pays ou provinces sont variables et ne peuvent être appliqués au contexte québécois, l'INESSS a développé un modèle économique (figure 4) afin d'évaluer l'efficacité des tests moléculaires (ThyroSeq v3 et Afirma GSC) pour préciser le diagnostic des nodules à cytologie indéterminée comparativement à la lobectomie. La perspective retenue est celle du système de soins de santé québécois. Il s'agit d'une analyse coût-utilité basée sur un modèle mixte (arbre de décision et modèle de Markov).

Figure 4 **Modèle d'analyse de l'efficacité**



L'arbre de décision proposé a pour objectif de modéliser l'issue clinique des patients selon l'approche diagnostique employée (lobectomie vs tests moléculaires). Le modèle de Markov avec des cycles annuels sur un horizon temporel de cinq ans a servi pour modéliser la prise en charge après la chirurgie (lobectomie) ou la surveillance des nodules moléculairement malins ou bénins. La période suivant la lobectomie est caractérisée par la présence ou l'absence de complications. Le modèle a inclus seulement l'hypothyroïdie comme complication, les autres étant considérées comme rares.

La surveillance a inclus le suivi des nodules moléculairement bénins (faux négatifs et vrais négatifs) et celui des nodules moléculairement suspects pour lesquels une surveillance a été jugée appropriée. Un suivi échographique annuel d'une durée maximale de cinq ans pour les nodules stables a été considéré selon l'algorithme de prise en charge des nodules thyroïdiens ([annexe B](#)).

2.9.2.1. Intrants cliniques et économiques

Les intrants cliniques et économiques considérés dans le modèle en lien avec les nodules thyroïdiens à cytologie indéterminée sont présentés dans le tableau 16. Les intrants cliniques (performance des tests, probabilités des événements, scores d'utilité) proviennent de la revue de la littérature scientifique. Les intrants de coûts ont été extraits des bases de données administratives du Québec (RAMQ et MSSS). Les prix des tests ont été obtenus auprès du MSSS.

Tableau 16 Intrants cliniques et économiques considérés dans le modèle

Description des variables	Valeur de base	Valeur pour l'analyse de sensibilité IC95 %	Distribution	Référence
Risque de cancer pour les nodules à cytologie indéterminée	0,25	0,10 - 0,40	Beta	Cibas et Ali, 2017
Sensibilité (ThyroSeq v3)	0,94	0,86 - 0,98	Beta	Steward <i>et al.</i> , 2019
Spécificité (ThyroSeq v3)	0,82	0,75 - 0,87	Beta	
Sensibilité (Afirma GSC)	0,91	0,79 - 0,98	Beta	Patel <i>et al.</i> , 2018
Spécificité (Afirma GSC)	0,68	0,60 - 0,76		
Résultat non concluant	0,10	0,05 - 0,15	Beta	Steward <i>et al.</i> , 2019
Probabilité de résection en cas de cytologie indéterminée	0,55	0,40 - 0,80	Beta	Bongiovanni <i>et al.</i> , 2012
Probabilité de résection en cas de résultat moléculairement suspect	0,77	0,54 - 0,97	Beta	Selon les taux rapportés dans la section 2.4 pour Afirma GEC, GSC, ThyroSeq v2.1 et 3
Probabilité de résection en cas de résultat moléculairement bénin	0,13	0 - 0,48	Beta	
Probabilité annuelle de diagnostic de cancer ultérieur si le résultat est moléculairement bénin	0,20	0,1 - 0,33	Beta	Balentine <i>et al.</i> , 2018
Risque d'hypothyroïdie associé à la lobectomie	0,22	0,19 - 0,27	Beta	Verloop <i>et al.</i> , 2012
Risque relatif de mortalité pour un cancer de la thyroïde non détecté	1,80	1,5 - 2,5	Log-normale	Société canadienne du cancer (SCC), 2019
Risque de décès dans la population générale (50-65 ans)	0,005	0,003 - 0,007	Beta	Statistique Canada. <i>Canada life tables</i>
Coûts (détails au tableau F-5 de l'annexe F)				
Coût de ThyroSeq v3 et Afirma GSC	5 385 \$	± 25 %	Gamma	INESSS*
Coût de la lobectomie	9 842 \$	-	Fixe	MSSS et RAMQ
Coût annuel associé à la surveillance d'un nodule bénin ou malin non réséqué	178,76 \$	-	Fixe	MSSS et RAMQ
Coût annuel associé au traitement et à la surveillance de l'hypothyroïdie	158,07 \$	-	Fixe	MSSS et RAMQ
Scores d'utilité†				
Utilité surveillance d'un nodule bénin/malin sans chirurgie	0,98	0,90 - 0,99	Beta	Li <i>et al.</i> , 2011
Utilité (hypothyroïdie postlobectomie)	0,95	0,877 - 0,991	Beta	Li <i>et al.</i> , 2011
Utilité postchirurgie sans complications	0,99	0,90 - 0,99	Beta	Li <i>et al.</i> , 2011
Décrément d'utilité lobectomie	-0,013	0,005 - 0,057-	Beta	Balentine <i>et al.</i> , 2018

* Approximation à partir de différentes sources.

† Comme les scores d'utilité sont plus élevés, ils ont été ajustés par rapport au niveau des scores de la population canadienne en excellente santé [Guertin *et al.*, 2018].

2.9.2.2. Résultats

Étant donné que l'efficacité des tests moléculaires ThyroSeq v3 et Afirma GSC n'a pas fait l'objet d'une comparaison directe, l'évaluation de l'efficacité de chaque test a été réalisée séparément, en comparaison avec la lobectomie. Les tableaux 17 et 18 présentent les résultats du scénario de base de l'analyse coût-utilité pour ThyroSeq v3 et Afirma GSC, respectivement.

En considérant un horizon temporel de cinq ans, l'utilisation de ThyroSeq v3 domine la lobectomie diagnostique. En effet, l'option ThyroSeq v3 serait légèrement moins coûteuse (5 972 \$ par personne vs 6 009 \$ pour la lobectomie) et légèrement plus efficace (4,381 QALY par personne vs 4,359 pour la lobectomie).

Tableau 17 Résultats du scénario de base de l'analyse coût-utilité de ThyroSeq v3 versus la lobectomie

Option	Coûts moyens totaux par personne	QALY par personne	Δ coûts par personne	Δ QALY par personne	RCUD
ThyroSeq v3	5 972 \$	4,381			
Statu quo (lobectomie)	6 009 \$	4,359	37 \$	-0,022	Dominée

QALY : *quality-adjusted life year*; RCUD : rapport coût-utilité différentiel.

Par ailleurs, l'utilisation d'Afirma GSC serait plus coûteuse que la lobectomie (6 287 \$ par personne vs 6 009 \$ pour la lobectomie) et légèrement plus efficace (4,371 QALY par personne vs 4,359 pour la lobectomie) avec un rapport coût-utilité différentiel de 22 667 \$ par QALY gagnée.

Tableau 18 Résultats du scénario de base de l'analyse coût-utilité d'Afirma GSC versus la lobectomie

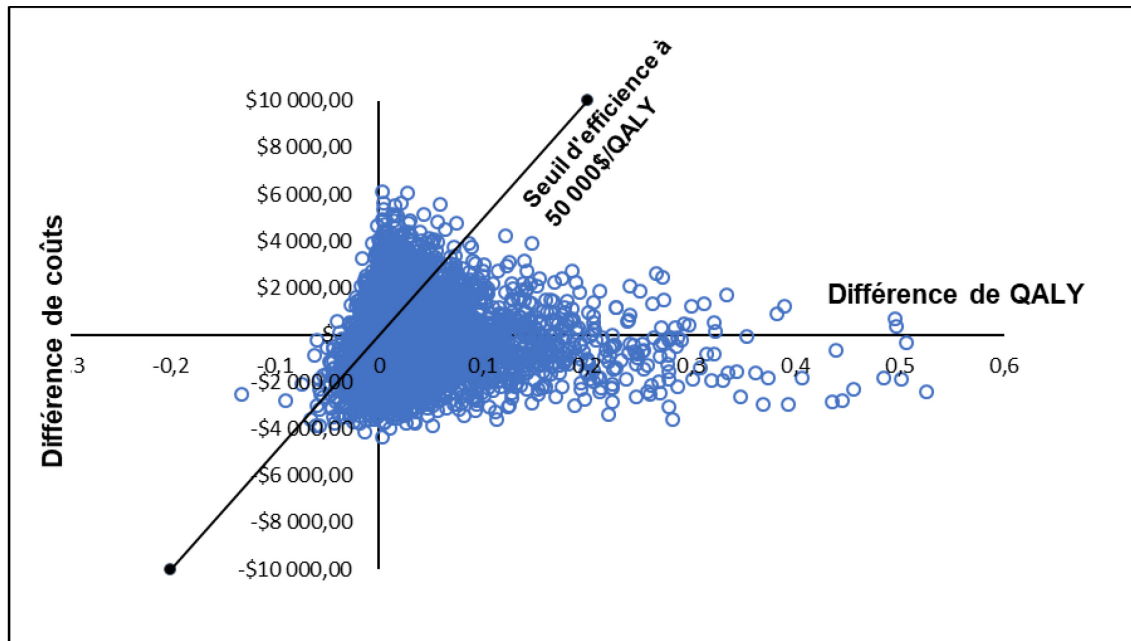
Option	Coûts moyens totaux par personne	QALY par personne	Δ coûts par personne	Δ QALY par personne	RCUD
Statu quo (lobectomie)	6 009 \$	4,359			
Afirma GSC	6 281 \$	4,371	272 \$	0,012	22 667 \$

QALY : *quality-adjusted life year*; RCUD : rapport coût-utilité différentiel.

2.9.2.3. Analyses de sensibilité

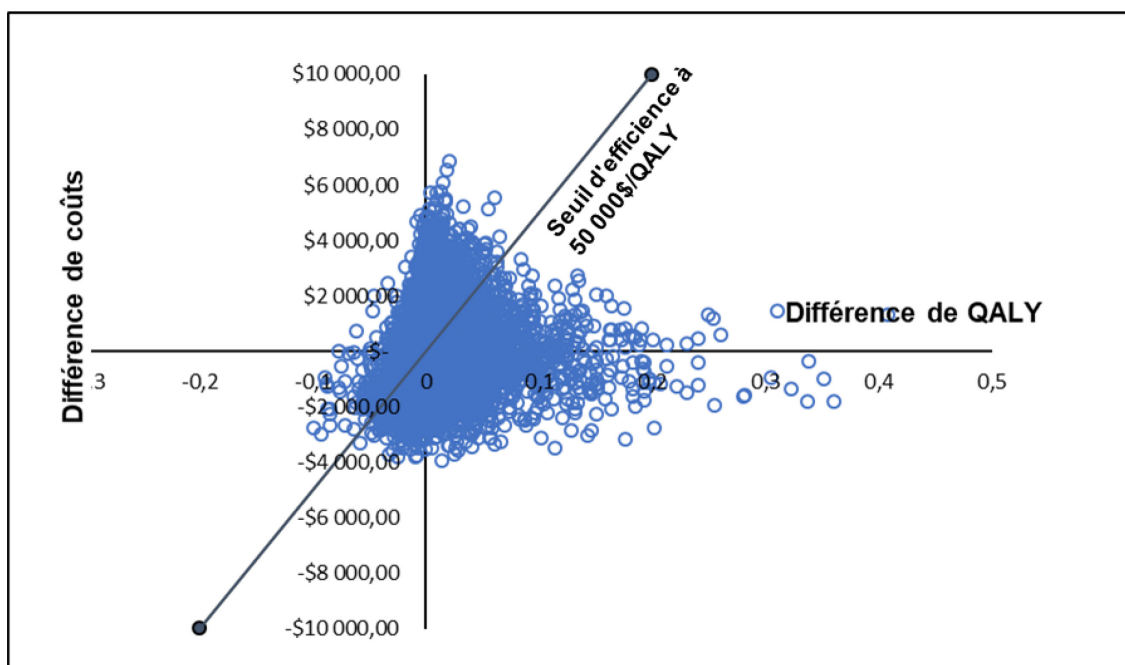
Les analyses de sensibilité probabilistes ont été réalisées par simulations de Monte Carlo (10 000 itérations) où tous les paramètres (tableau 16) varient simultanément selon leurs fonctions de distribution. Ces analyses ont montré qu'au seuil de 50 000 \$/QALY gagnée l'option ThyroSeq v3 présente une probabilité d'efficacité de 69 %; cette probabilité s'élève à 74,5 % lorsque le seuil de 100 000 \$/QALY gagnée est appliqué (figure 5).

Figure 5 Nuage de points ThyroSeq v3 vs lobectomie



Les analyses de sensibilité réalisées pour le test Afirma GSC ont par ailleurs révélé qu'aux seuils de 50 000 et 100 000 \$/QALY gagnée la probabilité d'efficacité est incertaine (52 % et 57 %, respectivement) (figure 6).

Figure 6 Nuage de points Afirma GSC vs lobectomie



EN BREF

La preuve de l'utilité clinique des tests moléculaires est incertaine. Toutefois, sur la base des hypothèses formulées, la probabilité de l'efficacité de cette approche pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée serait variable en comparaison avec la lobectomie. Aux seuils généralement acceptés :

- l'usage de ThyroSeq v3 serait probablement efficient;
- l'efficacité associée à l'utilisation du test Afirma GSC est incertaine.

2.9.3. Analyse d'impact budgétaire

L'analyse d'impact budgétaire prend en considération les coûts liés au remboursement potentiel des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée. Elle a été réalisée à partir des données épidémiologiques et médico-administratives (bases de données du MSSS et de la RAMQ) ainsi que selon l'information obtenue auprès des experts consultés. Les coûts sont projetés sur un horizon temporel de trois ans selon la perspective du système de soins de santé québécois comme payeur.

2.9.3.1. Population cible

Le nombre de patients qui présentent un nodule thyroïdien à cytologie indéterminée et qui pourraient bénéficier des tests moléculaires pour préciser le diagnostic a été évalué à 969 annuellement. Ce nombre a été obtenu en considérant les éléments suivants :

- 2 486 chirurgies thyroïdiennes ont été réalisées au Québec en 2018 (données provenant de la RAMQ),
- 43,2 % des nodules réséqués présentent une cytologie indéterminée [Bongiovanni *et al.*, 2012];
- 95 % des nodules à cytologie indéterminée réséqués le sont par lobectomie (opinion d'experts);
- 95 % des lobectomies sont à visée diagnostique (opinion d'experts).

Le tableau 19 présente la répartition des patients selon le scénario statu quo (lobectomie) et le nouveau scénario avec l'option ThyroSeq v3 ou Afirma GSC. Les détails sont présentés aux tableaux F-1 et F-2 de l'[annexe F](#).

Tableau 19 Répartition des patients selon le scénario

	AN 1	AN 2	AN3	TOTAL
Statu quo : Lobectomie				
Nombre de lobectomies	969	969	969	2 907
Nombre d'hypothyroïdies post-lobectomie	213	426	639	639
Nouveau scénario : Option ThyroSeq v3				
Nombre total de lobectomies	356	356	356	1 068
Nombre de nodules sous surveillance	613	1 226	1 839	1 839
Nombre d'hypothyroïdies post-lobectomie	78	156	234	234
Nouveau scénario : Option Afirma GSC				
Nombre total de lobectomies	416	416	416	1 248
Nombre de nodules sous surveillance	553	1 106	1 659	1 659
Nombre d'hypothyroïdies post-lobectomie	92	184	276	276

2.9.3.2. Résultats

Le tableau 20 présente les résultats du scénario de base pour l'analyse portant sur ThyroSeq v3. En considérant les intrants cliniques et économiques de la section sur l'efficacité (tableau 16), le coût annuel du nouveau scénario ThyroSeq v3 serait de 9,06 M\$, 9,09 M\$ et 9,13 M\$ pour chacune des 3 premières années, respectivement, pour un total d'environ 27,28 M\$ sur 3 ans.

Le remboursement public du test ThyroSeq v3 pour préciser le diagnostic des nodules à cytologie indéterminée pourrait permettre des économies de coûts évaluées à 408 491 \$, 320 124 \$ et 231 858 \$ pour chacune des 3 premières années, respectivement, pour un total de 960 473 \$.

Tableau 20 Impact budgétaire associé au remboursement de ThyroSeq v3

	AN 1	AN 2	AN 3	TOTAL
Statu quo : lobectomie				
Coûts des lobectomies*	9 027 517 \$	9 027 517 \$	9 027 517 \$	27 082 551 \$
Coûts associés à l'hypothyroïdie post-lobectomie†	33 706 \$	67 412 \$	101 117 \$	202 235 \$
Total	9 061 223 \$	9 094 929 \$	9 128 634 \$	27 284 786 \$
Nouveau scénario : option ThyroSeq v3				
Coûts des tests	5 219 366 \$	5 219 366 \$	5 219 366 \$	15 658 099 \$
Coûts des lobectomies*	3 311 293 \$	3 311 293 \$	3 311 293 \$	9 933 880 \$
Coûts associés à l'hypothyroïdie post-lobectomie†	109 709 \$	219 419 \$	329 128 \$	658 256 \$
Coûts associés à la surveillance des nodules‡	12 363 \$	24 727 \$	36 988 \$	74 078 \$
Total	8 652 732 \$	8 774 805 \$	8 896 776 \$	26 324 313 \$
Impact net	- 408 491 \$	- 320 124 \$	- 231 858 \$	- 960 473 \$

* Honoraires des médecins, coûts d'établissements (bloc opératoire, hospitalisation).

† Visites médicales, hormonothérapie de substitution, examens de laboratoire.

‡ Visites médicales et échographies de suivi selon l'algorithme de l'[annexe B](#).

Le tableau 21 présente les résultats du scénario de base pour l'analyse d'Afirma GSC. En considérant les intrants cliniques et économiques de la section sur l'efficience (tableau 16), le coût annuel du nouveau scénario Afirma GSC serait de 9,21 M\$, 9,32 M\$ et 9,43 M\$ pour chacune des 3 premières années, respectivement, pour un total d'environ 27,96 M\$ sur 3 ans.

Le remboursement public d'Afirma GSC pour préciser le diagnostic des nodules à cytologie indéterminée pourrait engendrer des coûts additionnels évalués à 146 118 \$, 225 777 \$ et 305 435 \$ pour chacune des 3 premières années, respectivement, pour un total de 677 330 \$.

Tableau 21 Impact budgétaire associé au remboursement d’Afirma GSC

	AN 1	AN 2	AN 3	TOTAL
Statu quo : lobectomie				
Coûts des lobectomies*	9 027 517 \$	9 027 517 \$	9 027 517 \$	27 082 551 \$
Coûts associés à l’hypothyroïdie post-lobectomie†	33 706 \$	67 412 \$	101 117 \$	202 235 \$
Total	9 061 223 \$	9 094 929 \$	9 128 634 \$	27 284 786 \$
Nouveau scénario : option Afirma GSC				
Coûts des tests	5 219 366 \$	5 219 366 \$	5 219 366 \$	15 658 099 \$
Coûts des lobectomies*	3 874 610 \$	3 874 610 \$	3 874 610 \$	11 623 831 \$
Coûts associés à l’hypothyroïdie post-lobectomie†	98 898 \$	197 795 \$	296 693 \$	593 386 \$
Coûts associés à la surveillance des nodules‡	14 467 \$	28 933 \$	43 400 \$	86 799 \$
Total	9 207 341 \$	9 320 705 \$	9 434 070 \$	27 962 116 \$
Impact net	146 118 \$	225 777 \$	305 435 \$	677 330 \$

* Honoraires des médecins, coûts d’établissements (bloc opératoire, hospitalisation).

† Visites médicales, hormonothérapie de substitution, examens de laboratoire.

‡ Visites médicales et échographies de suivi selon l’algorithme de l’annexe B.

2.9.3.3. Analyse des scénarios

Les analyses de scénarios ([tableaux F-3 et F-4](#) de l’annexe F) ont montré que le remboursement de ThyroSeq v3 pour préciser le diagnostic des nodules à cytologie indéterminée pourrait engendrer des économies pouvant varier entre 82 000 \$ (lorsque la spécificité du ThyroSeq v3 est fixée à 75 %) et 4,9 M\$ (lorsque son coût unitaire est diminué de 25 %) sur 3 ans. Cependant, le remboursement du test pourrait engendrer des coûts dans les circonstances suivantes (si tous les autres paramètres demeurent constants) :

- le risque de malignité pour un nodule à cytologie indéterminée est supérieur à 32,5 %;
- la proportion de nodules moléculairement suspects réséqués est supérieure à 87 %;
- la proportion de nodules moléculairement bénins réséqués dépasse 18,8 %;
- le coût unitaire du test est supérieur à 5 715 \$.

Les analyses des scénarios ont montré que le remboursement d’Afirma GSC pour préciser le diagnostic des nodules à cytologie indéterminée pourrait engendrer des coûts additionnels pouvant varier entre 173 000 \$ (sensibilité du test égale à 79 %) et 5,6 M\$ (lorsque la proportion de nodules moléculairement bénins réséqués est de 48 %) sur 3 ans. Le remboursement d’Afirma GSC pourrait diminuer les coûts dans les circonstances suivantes (si tous les autres paramètres demeurent constants) :

- le risque de malignité pour un nodule à cytologie indéterminée est inférieur à 18,2 %;
- la spécificité de l'analyse Afirma GSC est supérieure à 73,3 %;
- la proportion de nodules moléculairement suspects réséqués est inférieure à 71,5 %;
- la proportion de nodules moléculairement bénins réséqués est inférieure à 8,2 %;
- le coût unitaire du test est inférieur à 5 152 \$.

EN BREF

Le remboursement public de l'analyse ThyroSeq v3 pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée pourrait diminuer les coûts de 0,960 M\$ (0,173 - 5,6 M\$) sur 3 ans. Le remboursement du test pourrait augmenter les coûts si :

- le risque de malignité pour un nodule à cytologie indéterminée est supérieur à 32,5 %;
- la proportion de nodules moléculairement suspects réséqués est supérieure à 87 %;
- la proportion de nodules moléculairement bénins réséqués dépasse 18,8 %;
- le coût unitaire du test est supérieur à 5 715 \$.

Le remboursement public d'Afirma GSC pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée pourrait augmenter les coûts de 0,677 M\$ (0,083 M\$ - 5,5 M\$) sur 3 ans. Cependant, le remboursement du test pourrait diminuer les coûts si :

- le risque de malignité pour un nodule à cytologie indéterminée est inférieur à 18,2 %;
- la spécificité de l'analyse Afirma GSC est supérieure à 73,3 %;
- la proportion de nodules moléculairement suspects réséqués est inférieure à 71,5 %;
- la proportion de nodules moléculairement bénins réséqués est inférieure à 8,2 %;
- le coût unitaire du test est inférieur à 5 152 \$.

DISCUSSION

Le nodule thyroïdien est une pathologie relativement fréquente, mais elle est maligne dans une minorité de cas. La biopsie par aspiration à l'aiguille fine est communément pratiquée dans le but de déterminer la nature du nodule. Depuis son introduction en 2009, le système de classification Bethesda a grandement facilité les décisions de prise en charge de la plupart des nodules thyroïdiens soumis à l'évaluation cytologique [Cibas et Ali, 2017; Cibas et Ali, 2009]. Dans 20 à 30 % des cas cependant, le résultat de la cytologie est indéterminé et la conduite optimale incertaine. Dans ces circonstances, le seul moyen d'établir un diagnostic définitif consiste à procéder à une résection chirurgicale. Toutefois, la majorité des nodules à cytologie indéterminée sont véritablement bénins et, de ce fait, leur résection est jugée superflue. À cela s'ajoute un risque non négligeable pour le patient de développer une hypothyroïdie post-opératoire, nécessitant la prise d'une hormonothérapie de substitution.

L'intérêt manifesté à l'égard des tests moléculaires résulte de leur présumée capacité à prédire le diagnostic histologique bénin des nodules à cytologie indéterminée, sans recours à la chirurgie. Les tests moléculaires qui affichent cette fonction (exclusion de la malignité) ont été évalués.

Performance des tests moléculaires disponibles

Les analyses moléculaires présentement disponibles, Afirma GSC et ThyroSeq v3, ont démontré une sensibilité et une valeur prédictive négative élevées dans le contexte des études de validation clinique originales [Steward *et al.*, 2019; Patel *et al.*, 2018]. Malgré l'absence d'études indépendantes confirmant leur performance pour prédire le diagnostic histologique bénin des nodules à cytologie indéterminée, les cliniciens (majoritairement américains) ont rapidement intégré ces outils dans leur pratique. Le principal risque associé à l'adoption précoce des tests moléculaires concerne la validité externe des conclusions issues de l'étude originale. La généralisation des caractéristiques de la performance d'un test dépend du caractère représentatif de la population évaluée par rapport à celle ciblée pour l'analyse dans la pratique courante. Selon les experts consultés, les nodules de taille inférieure à 1 cm de même que ceux de plus de 4 cm ne seraient pas soumis à l'évaluation moléculaire en contexte québécois, alors qu'ils étaient inclus lors de la validation clinique. Aussi, les caractéristiques échographiques des nodules, ou la catégorie de risque associée à celles-ci, n'ont pas été décrites par les auteurs des études cliniques mais pourraient avoir exercé un effet non négligeable sur les résultats observés. Malgré tout, les experts interrogés étaient d'avis qu'une sélection des patients basée sur des critères restreints pourrait mener à des caractéristiques de performance similaires à celles observées dans les études de validation clinique originales. Ces critères incluent les nodules à cytologie de catégorie Bethesda III (après une deuxième BAAF) ou IV, dont la taille est comprise entre 1 et 4 cm et qui ne présentent pas de facteur de risque clinique ou échographique élevé. Les patients qui répondent d'emblée à une indication pour la chirurgie, de même que ceux qui manifestent une préférence pour la chirurgie ou l'observation, ne devraient pas se voir

offrir une évaluation moléculaire. Son utilisation devrait être restreinte aux patients chez qui le résultat du test est susceptible d'influer sur la prise en charge.

Prévalence locale de la maladie parmi les nodules à cytologie indéterminée

La sensibilité et la spécificité sont des propriétés intrinsèques au test diagnostique : une sensibilité de 90 % signifie que, peu importe la population évaluée, 90 % des nodules malins auront un résultat moléculairement suspect. La véritable performance d'un test à l'extérieur de la cohorte de validation, mesurée en termes de valeur prédictive, dépend de la prévalence locale de la maladie parmi les nodules à cytologie indéterminée [Ferris *et al.*, 2015]. La valeur prédictive négative est donc modifiée chaque fois que le test est réalisé auprès d'une population qui présente une différence dans la prévalence de la maladie. Les études de validation clinique pour ThyroSeq v3 et Afirma GSC ont été réalisées sur 10 et 43 sites de prélèvement, respectivement. Globalement, la prévalence de la maladie a été de 28 % pour ThyroSeq v3 et de 24 % pour Afirma GSC. Cependant, parmi les sites qui ont fourni plus d'un échantillon, la prévalence locale de la maladie a varié jusqu'à 60 % dans l'étude pour ThyroSeq v3 et jusqu'à 56 % dans celle pour Afirma GSC ([tableau E-1](#)). Selon le théorème de Bayes et les critères reconnus par le NCCN et l'ATA, l'utilisation de ThyroSeq v3 et d'Afirma GSC serait jugée sécuritaire ($VPN \geq 95\%$) jusqu'à l'atteinte d'une prévalence maximale de la maladie de 40 % et 29 %, respectivement. On trouve d'ailleurs, dans le [rapport d'un résultat bénin](#) avec l'analyse ThyroSeq v3, une note du fabricant selon laquelle le risque résiduel de malignité observé en présence d'un tel résultat (3 %) est applicable à un contexte clinique dans lequel le risque prétest de maladie parmi les nodules à cytologie indéterminée est inférieur à 40 %.

Selon la méta-analyse de Bongiovanni, le recours à la catégorie « indéterminée » pour caractériser le résultat de la cytologie est très variable d'une étude à l'autre (BIII : 0,8-27,2 %; BIV : 1,2-25,3 %) [Bongiovanni *et al.*, 2012]. Une certaine variation dans l'interprétation du résultat de la cytologie par différents cytopathologistes a également été documentée [Padmanabhan *et al.*, 2017]. La capacité d'attribuer le diagnostic cytologique approprié influe sur la prévalence de la maladie associée à chaque catégorie et, par conséquent, sur la valeur prédictive des tests moléculaires. Les données contextuelles québécoises ont relevé la présence d'écarts importants quant au volume annuel de biopsies par aspiration à l'aiguille fine compilé pour chaque établissement de santé. Ces disparités pourraient contribuer à accentuer la variation observée quant à la propension des cytopathologistes à utiliser la catégorie cytologique indéterminée. Par conséquent, la prévalence de la maladie dans cette catégorie ainsi que la performance des tests sont susceptibles de varier grandement d'un établissement à l'autre.

Même s'il est généralement admis qu'une augmentation de la prévalence de la maladie pourrait avoir un effet néfaste sur la performance locale des tests pour l'identification des nodules bénins, les experts consultés se sont dits peu préoccupés par cette variable et son impact sur le risque résiduel de maladie lorsque le résultat est moléculairement bénin (augmentation du taux de faux négatifs). Selon eux, le suivi clinique/échographique réservé aux nodules moléculairement bénins, couplé à la progression lente des lésions

malignes de stade précoce, permettent d'anticiper que les nodules faussement négatifs qui auront échappé à la détection moléculaire seront éventuellement décelés sans que cela porte un important préjudice aux patients concernés. Dans un milieu où la prévalence locale de la maladie parmi les nodules à cytologie indéterminée atteindrait 50 %, le risque résiduel pour un patient qui présente un nodule à cytologie indéterminée moléculairement bénin (7 % avec ThyroSeq v3 et 12 % avec Afirma GSC) serait inférieur au risque initial que présente le nodule dans ce milieu. Néanmoins, le suivi des nodules moléculairement bénins devrait être fait sous étroite surveillance afin d'assurer le repérage précoce des nodules faussement négatifs qui pourraient évoluer vers des cancers agressifs. Le cas des patients perdus de vue au suivi clinique représente une préoccupation dans ce contexte.

Utilité clinique des tests moléculaires

Le taux historique de chirurgies pratiquées dans les cas de nodules à cytologie indéterminée se situe autour de 55 % [Bongiovanni *et al.*, 2012]. Selon les experts consultés, un pourcentage similaire de nodules à cytologie indéterminée auraient été réséqués au Québec au cours des dernières années. L'utilité clinique des tests moléculaires devrait se traduire par une réduction du nombre de chirurgies thyroïdiennes parmi les nodules à cytologie indéterminée, en comparaison avec la pratique actuelle. Le groupe comparateur représentant la pratique actuelle devrait être évalué au cours de la même période que le groupe expérimental, mais en l'absence de test moléculaire. À défaut de répartition aléatoire (*randomisation*), les analyses devraient être ajustées pour les variables liées aux patients et à la maladie. Cependant, aucune des études sélectionnées ne répondait à ces critères. De plus, comme la durée médiane du suivi était rarement précisée par les auteurs, il était généralement difficile de savoir si celle-ci avait été suffisamment longue pour permettre l'identification de tous les nodules qui avaient été réséqués à la suite de l'évaluation moléculaire. Par ailleurs, une évaluation adéquate de l'utilité devrait reposer sur l'emploi sélectif du test : en l'absence d'indication claire pour une chirurgie chez les patients qui n'ont pas de préférence marquée pour la chirurgie ou l'observation. Dans plusieurs études, les critères qui guident l'usage du test étaient inconnus ou absents (usage systématique); lorsque l'usage était sélectif, les critères appliqués pour la sélection n'étaient que rarement énoncés.

L'utilité clinique de l'analyse ThyroSeq v3 a été évaluée au moyen d'une seule étude portant sur un nombre restreint de nodules à cytologie indéterminée (n = 48). Le taux de chirurgies rapporté pour l'ensemble des nodules a été de 44 %. Un biais lié à la sélection des cas à risque élevé est possiblement responsable du taux élevé de chirurgies observé dans cette étude, et souligne l'efficacité limitée de ce type d'analyse (*rule-out test*) pour prévenir les chirurgies inutiles parmi les nodules à risque élevé. La proportion de nodules TI-RADS 5 moléculairement bénins n'a pas été rapportée. L'histologie des nodules moléculairement bénins n'était connue que pour 2 nodules sur 28 (7 %). Considérant le nombre de nodules histologiquement malins réséqués dans cette étude (19/48), la prévalence locale de la maladie parmi les nodules à cytologie indéterminée était supérieure au seuil de 40 % défini par [ThyroSeq](#) pour éviter la chirurgie de façon sécuritaire en cas de résultat moléculairement bénin. La valeur prédictive négative et le

risque résiduel de maladie des nodules moléculairement bénins qui ne sont pas réséqués n'ont pas été validés dans ce contexte.

Parmi les autres études sélectionnées pour évaluer l'utilité clinique des tests moléculaires, le taux global de chirurgies rapporté après avoir combiné tous les nodules par type d'analyse a été de 51 % avec Afirma GEC, de 25 % avec Afirma GSC et de 44 % avec ThyroSeq v2.1. Si l'avantage qui résulte de l'utilisation des versions antérieures paraît limité, l'emploi du test Afirma GSC pourrait s'avérer profitable, à condition bien sûr que la sélection initiale des patients ait été adéquate et que le taux de chirurgies observé à la suite de ce test demeure stable à plus long terme. Étant donné la mise en application récente de l'analyse, des études au suivi prolongé seront nécessaires pour confirmer son utilité clinique. En l'absence d'un comparateur adéquat, il est aussi difficile de spéculer quant à l'ampleur du gain généré. Une comparaison indirecte pourrait être faite en employant une valeur historique comme celle rapportée en méta-analyse (55 % des nodules à cytologie indéterminée sont réséqués) [Bongiovanni *et al.*, 2012]. Toutefois, ce taux pourrait s'appliquer à une population beaucoup plus large que celle sélectionnée pour l'évaluation moléculaire, puisqu'il inclut aussi les nodules à cytologie indéterminée réséqués en raison de la présence de facteurs de risque cliniques ou échographiques élevé. Par ailleurs, aucune étude clinique en cours évaluant l'utilité clinique des analyses présentement disponibles en comparaison avec la pratique actuelle n'a été repérée. Un essai clinique *randomisé* est actuellement mené ([NCT02681328](#)) afin de déterminer lequel des tests, ThyroSeq v3 ou Afirma GSC, permet à un plus grand nombre de patients d'éviter une chirurgie inutile. Les résultats de la phase initiale de cette étude réalisée avec les versions antérieures (Afirma GEC et ThyroSeq v2.1) ont été présentés dans les sections 2.4.1 et 2.4.3 de cet avis. Rappelons que, dans cette étude, l'analyse moléculaire a été prescrite de façon systématique lorsque le résultat de la cytologie était indéterminé. La date de complétion primaire de la seconde phase de l'étude est attendue pour août 2021.

Le taux de résultats moléculairement bénins (BCR) représente la population cible qui pourrait éviter une chirurgie diagnostique. Étant donné qu'environ 75 % des nodules à cytologie indéterminée réséqués (incluant ceux qui présentent des facteurs de risque élevé) sont histologiquement bénins, un test devrait optimalement avoir un BCR qui tend vers cette valeur. Parmi les études d'utilité clinique retenues, la valeur globale du BCR a été de 44 % pour Afirma GEC, de 68 % pour Afirma GSC et de 74 % pour ThyroSeq v2.1. Ces données, couplées aux taux de chirurgies observés après une évaluation avec chaque test suggèrent que certains manquent de spécificité (Afirma GEC) ou présentent un taux élevé de nodules moléculairement bénins réséqués.

Le taux de chirurgies observé parmi les nodules moléculairement bénins dépend de plusieurs facteurs, dont les critères de sélection appliqués pour soumettre les nodules à l'évaluation moléculaire. Tel que mentionné précédemment, l'emploi judicieux des tests moléculaires requiert l'exclusion des patients qui présentent d'emblée une indication claire ou une préférence marquée pour la chirurgie. Le non-respect de ces conditions comporte le risque de provoquer une augmentation biaisée du nombre de résections. Par exemple, l'inclusion des nodules qui présentent des facteurs de risque élevé pourrait

augmenter le taux de faux négatifs et accroître ainsi le nombre de chirurgies lorsque les nodules faussement négatifs auront été décelés et réséqués, comme il se doit. Par ailleurs, l'inclusion des patients qui préfèrent la chirurgie à l'observation conduira inévitablement à une augmentation erronée du nombre de chirurgies parce que, même en présence d'un résultat moléculairement bénin, ces patients souhaiteront que leur nodule soit réséqué. Le degré de confiance des cliniciens dans le résultat du test est également susceptible d'affecter la prise en charge subséquente. Lorsque l'ensemble des nodules ont été considérés pour chaque test, le taux de nodules moléculairement bénins réséqués a été de 14 % avec Afirma GEC, de 6 % avec Afirma GSC et de 33 % avec ThyroSeq v2.1. Parmi ceux-ci, 80 %, 81 % et 92 % étaient histologiquement bénins et n'auraient donc pas dû être réséqués.

Analogie entre nodules à cytologie indéterminée moléculairement bénins et nodules à cytologie bénigne

Le risque de malignité des nodules à cytologie bénigne est considéré comme suffisamment faible (< 5 %) pour qu'une chirurgie soit évitée chez la majorité des patients. Par analogie, un test capable d'identifier les nodules bénins avec un risque résiduel similaire devrait aussi permettre aux nodules évalués au moyen de cette méthode d'éviter une chirurgie. Les données présentées par Angell et ses collaborateurs suggèrent que les nodules moléculairement bénins ont un comportement clinique semblable à celui des nodules à cytologie bénigne, justifiant ainsi une prise en charge similaire [Angell *et al.*, 2015]. On peut donc anticiper une proportion comparable de nodules réséqués dans chaque groupe. Historiquement, le taux de résection parmi les nodules à cytologie bénigne a été évalué à 10,4 % [Bongiovanni *et al.*, 2012]. Peu d'études ont réalisé une comparaison directe entre les nodules à cytologie indéterminée moléculairement bénins et ceux à cytologie bénigne; les deux seules repérées concernent la version GEC d'Afirma. Le suivi était généralement de courte durée, compte tenu de la progression lente des lésions malignes de stade précoce, mais il était néanmoins similaire dans les groupes comparés. Dans l'une des études, le taux rapporté de chirurgies pour les nodules GEC bénins a été semblable à celui observé pour les nodules à cytologie bénigne (11,4 % vs 10,1 %; $p = 0,594$). Dans l'autre, une différence statistiquement significative a été rapportée (13,8 % pour les nodules GEC bénins vs 0,9 % pour les nodules à cytologie bénigne; $p < 0,0001$). Par contre, la valeur compilée pour les nodules à cytologie bénigne (0,9 %) était considérablement inférieure à celle publiée par Bongiovanni et ses collaborateurs (10,4 %). De plus, la période d'évaluation était différente (2001-2010) de celle des nodules à cytologie indéterminée GEC bénins (2010-2014). Même si aucune étude n'a comparé la proportion de nodules réséqués parmi les nodules à cytologie bénigne à celle des nodules moléculairement bénins avec les versions présentement disponibles (Afirma GSC ou ThyroSeq v3), il apparaît probable que ces versions aux caractéristiques de performance améliorées conduisent minimalement à des taux de chirurgie similaires. Dans un contexte de données probantes limitées, nous ne pouvons toutefois pas en être certains.

L'examen des données issues d'études qui ont évalué l'utilité clinique des analyses moléculaires souligne aussi l'incertitude associée au risque résiduel de maladie parmi les nodules moléculairement bénins non réséqués. Le théorème de Bayes permet habituellement d'estimer le taux de faux négatifs attendu à partir de la prévalence locale de la maladie parmi les nodules à cytologie indéterminée. Mais, dans un contexte où l'utilisation des tests moléculaires contribue à limiter le nombre de nodules réséqués, la valeur de la prévalence locale de la maladie est désormais inconnue. Les approches élaborées par les auteurs dans le but d'estimer le risque résiduel de malignité à la suite d'un résultat moléculairement bénin sont imprécises, soit parce que les nodules non réséqués ont été exclus de l'équation ou parce que tous les nodules non réséqués ont été jugés d'emblée histologiquement bénins. Par ailleurs, aucune différence significative n'a été rapportée concernant le risque résiduel des nodules à cytologie indéterminée GEC bénins en comparaison avec celui associé aux nodules à cytologie bénigne après un suivi médian comparable de 13 et 14 mois (12,5 % vs 33,3 %; $p = 0,16$). Par contre, le nombre de nodules GEC bénins était limité ($n = 58$), la période d'évaluation des deux groupes était différente et la proportion de cas malins parmi les nodules à cytologie bénigne était considérablement supérieure à celle rapportée en méta-analyse pour 15 104 nodules à cytologie bénigne (33,3 % vs 3,7 %) [Bongiovanni *et al.*, 2012].

Perspective des patients

La perspective des patients a fait l'objet de deux publications. Selon les données rapportées, les patients ont préféré l'observation à la chirurgie dans la mesure où le risque de malignité demeurerait inférieur à un seuil moyen de 39 %. Cette valeur est comparable à la borne supérieure qui définit le risque de maladie pour l'ensemble des nodules à cytologie indéterminée (de 10 à 40 %), mais supérieure au risque résiduel observé (environ 5 %, selon la prévalence locale de la maladie) à la suite d'un résultat moléculairement bénin [Cibas et Ali, 2017]. Les patients ont aussi manifesté une préférence marquée pour l'évaluation moléculaire en comparaison avec la prise en charge standard des nodules à cytologie indéterminée. L'analyse de l'impact des tests moléculaires sur la qualité de vie des patients a révélé qu'un résultat bénin au test moléculaire serait aussi rassurant pour les patients qu'un résultat de cytologie bénigne. D'autres études seront nécessaires pour documenter la perspective des patients quant à la prise en charge subséquente des nodules à cytologie indéterminée moléculairement bénins. Bien que la lobectomie diagnostique présente certains inconvénients, elle possède néanmoins l'avantage de procurer un diagnostic histologique définitif. L'impact d'un suivi clinique/échographique prolongé sur la qualité de vie des patients qui présentent un nodule moléculairement bénin est encore mal connu.

Considérations économiques du remboursement des analyses moléculaires

Étant donné que les résultats des études économiques publiées sont variables et ne peuvent être appliqués au contexte québécois, l'INESSS a réalisé une analyse de l'efficacité et de l'impact budgétaire des tests moléculaires (ThyroSeq v3 et Afirma GSC) pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée. Comme la performance de chaque test n'a pas fait l'objet d'une comparaison directe sur les mêmes

nodules, l'évaluation de l'efficacité a été réalisée séparément, en comparaison avec la lobectomie.

Les résultats obtenus montrent que, sur la base des hypothèses formulées, lesquelles sont assorties d'incertitude, la probabilité de l'efficacité de l'approche moléculaire serait variable en comparaison avec la lobectomie. Au seuil de 50 000 \$/QALY gagnée, l'emploi de ThyroSeq v3 serait probablement efficace (69 %) alors que l'efficacité associée à l'utilisation du test Afirma GSC reste incertaine (52 %).

L'impact budgétaire associé au remboursement public des tests moléculaires serait aussi variable selon les scénarios envisagés. D'après le scénario de base, le remboursement de l'analyse ThyroSeq v3 pourrait diminuer les coûts de 0,960 M\$ sur 3 ans, comparativement à la lobectomie diagnostique. Par ailleurs, le remboursement d'Afirma GSC pourrait quant à lui augmenter les coûts de 0,677 M\$ sur 3 ans comparativement à la lobectomie diagnostique. L'analyse de scénarios a révélé que les résultats obtenus pour le scénario de base étaient sensibles à plusieurs variables incluant la prévalence de la maladie dans la catégorie indéterminée, la proportion de nodules réséqués selon le résultat moléculaire, la spécificité de l'analyse et son coût unitaire.

RÉSUMÉ DES DÉLIBÉRATIONS ET RECOMMANDATIONS

Principaux constats de l'évaluation

L'analyse des données scientifiques, contextuelles et expérientielles a permis de formuler les constats généraux suivants :

- Il existe un besoin clinique non comblé pour réduire, de façon non invasive, l'incertitude diagnostique associée aux nodules thyroïdiens à cytologie indéterminée.
- L'évaluation moléculaire a été développée pour faciliter l'identification des nodules à cytologie indéterminée qui s'avèrent histologiquement bénins. Cette approche permettrait de réduire le nombre de chirurgies diagnostiques pratiquées et éviterait à certains patients la prise d'hormones de substitution liée à l'hypothyroïdie post-opératoire.
- Trois tests sont présentement offerts par des laboratoires américains (analyses réalisées hors Québec) : ThyroSeq v3, Afirma GSC et ThyGeNext/ThyraMIR⁷. L'offre est livrée à un coût élevé et sans nécessité d'approbation réglementaire du produit/service par la Food and Drug Administration, puisque l'analyse est réalisée par le fabricant.
- ThyroSeq v3 et Afirma GSC ont une sensibilité (> 90 %) et une valeur prédictive négative (> 95 %) élevée. Leur performance n'a toutefois pas été confirmée par des études indépendantes. Par ailleurs, les données d'utilité clinique publiées pour ces versions sont limitées. Par conséquent, le niveau de preuve associé à la démonstration de la validité et de l'utilité clinique est faible et un doute persiste quant à leur performance et leurs impacts (cliniques et décisionnels) dans la population ciblée.
- Le risque résiduel de maladie à la suite d'un résultat moléculairement bénin est incertain, en raison de la variation de la prévalence de la maladie parmi les nodules à cytologie indéterminée dans chaque établissement de santé. Le cas des patients perdus de vue au suivi clinique représente une préoccupation dans ce contexte.
- À défaut d'un résultat moléculairement bénin, les nodules sélectionnés pour l'évaluation moléculaire devraient faire l'objet d'une lobectomie diagnostique.
- L'efficacité de ThyroSeq v3 et d'Afirma GSC comparativement à la lobectomie diagnostique serait variable selon les hypothèses postulées. Aux seuils d'efficacité généralement acceptés, l'utilisation de ThyroSeq v3 serait probablement efficace. Sur la base des données disponibles, la probabilité de l'efficacité associée à l'emploi du test Afirma GSC serait incertaine.

⁷ L'évaluation de la validité clinique de ce test est en cours.

- L'impact budgétaire lié au remboursement public des tests moléculaires serait également variable selon les scénarios envisagés. D'après le scénario de base, le remboursement de l'analyse ThyroSeq v3 pourrait diminuer les coûts comparativement à la lobectomie diagnostique, alors que le remboursement du test Afirma GSC pourrait les augmenter. Plusieurs variables, incluant la prévalence de la maladie dans la catégorie indéterminée, la proportion de nodules réséqués selon le résultat moléculaire, la spécificité de l'analyse et son coût unitaire sont susceptibles de faire varier les résultats du scénario de base.

Résumé des délibérations

Malgré l'incertitude relative à la validité et à l'utilité clinique de ces tests, les membres du comité délibératif (CEC santé et CSABM) en ont reconnu le caractère prometteur. Les membres sont d'avis que la couverture publique de ces tests pourrait constituer une décision juste et raisonnable, mais uniquement en présence de modalités qui tiendraient compte du contexte d'incertitude.

Motifs de la position :

- l'emploi de ce type de test, en remplacement de la lobectomie diagnostique, permettrait au patient d'éviter une chirurgie comportant un risque d'hypothyroïdie post-opératoire et libérerait du temps opératoire pour ce secteur d'activité ainsi que pour d'autres spécialités;
- un résultat moléculairement bénin pourrait constituer un facteur rassurant pour le patient, en comparaison avec le risque initial de maladie associé aux nodules à cytologie indéterminée;
- la progression lente des lésions malignes de stade précoce, et le suivi réservé aux nodules moléculairement bénins, permettraient d'envisager une prise en charge sécuritaire, même en cas de résultat faussement négatif;
- une économie de coûts est envisageable pour le diagnostic et la prise en charge des nodules à cytologie indéterminée, si l'efficacité anticipée de l'approche moléculaire est avérée;
- certains membres ont exprimé un malaise concernant le remboursement public de ce type d'analyse en raison de la difficulté rencontrée pour reconnaître la valeur de l'approche moléculaire en l'absence d'études de validation clinique non financées par le fabricant;
- des études de validation clinique indépendantes du fabricant ne sont pas attendues à court terme;
- il n'y a pas d'étude confirmant l'utilité clinique de ces tests en comparaison avec la pratique standard.

Recommandations

L'Institut national d'excellence en santé et en services sociaux (INESSS) reconnaît le caractère prometteur des tests ThyroSeq v3 et Afirma GSC pour préciser le diagnostic des nodules thyroïdiens dans certains cas de cytologie indéterminée. Toutefois, compte tenu de l'incertitude liée à la validité et à l'utilité clinique de ces tests et de l'effet de certaines variables sur les coûts de l'approche moléculaire, l'INESSS est d'avis que la couverture publique de ces tests pourrait constituer une décision juste et raisonnable si les conditions suivantes sont respectées :

- L'utilisation devrait être restreinte aux patients chez qui le résultat du test est susceptible d'influer sur la prise en charge.

Les nodules admissibles devraient avoir les caractéristiques suivantes :

- nodule à cytologie de catégorie Bethesda III (après une 2^e BAAF) ou Bethesda IV; et
- dont la taille est comprise entre 1 et 4 cm; et
- qui présentent des facteurs de risque cliniques ou échographiques intermédiaires (équivoques ou discordants).

Les patients pour lesquels une chirurgie est indiquée (facteurs de risque élevé), ceux pour qui la lobectomie diagnostique n'est pas envisagée en raison d'un faible risque de malignité, de même que ceux qui manifestent une préférence pour la chirurgie ou l'observation, ne devraient pas se voir offrir une évaluation moléculaire.

- La mise en œuvre préalable d'une étape de validation clinique en contexte québécois⁸.
- Une entente est conclue avec le fabricant pour partager le risque financier.

Autres éléments :

- Le patient doit être informé du risque résiduel de malignité associé à la décision de ne pas réséquer un nodule à cytologie indéterminée moléculairement bénin avant d'opter pour une analyse moléculaire.
- Les patients dont le nodule est moléculairement bénin devraient faire l'objet d'un suivi tous les 12 à 24 mois durant 4 à 5 ans afin d'assurer l'identification précoce des nodules faussement négatifs qui pourraient évoluer vers des cancers agressifs.

⁸ La confirmation indépendante de la validité clinique implique 1) que tous les nodules à cytologie indéterminée soumis au fabricant (selon les critères d'admissibilité mentionnés plus haut) soient aussi réséqués (dans le contexte de validation) pour que la performance de l'approche moléculaire soit établie sur la base du diagnostic histopathologique et 2) que les résultats issus de cette démonstration soient minimalement comparables à ceux générés dans le cadre de l'étude de validation clinique originale.

- L'INESSS suggère la réalisation d'une collecte de données prospectives afin de compiler l'information relative 1) aux caractéristiques de base des nodules sélectionnés pour l'évaluation moléculaire (Bethesda III ou IV, taille, caractéristiques échographiques), 2) au choix (ThyroSeq v3 ou Afirma GSC) et au résultat de l'analyse (bénin ou suspect) et 3) à la décision de prise en charge subséquente (observation ou chirurgie). Les données recueillies pourront ensuite faire l'objet d'une évaluation au terme de laquelle la révision des critères d'admissibilité ou un guide d'usage optimal pourrait être proposé.

RÉFÉRENCES

- Ablordeppey KK, Timmaraju VA, Song-Yang JW, Yaqoob S, Narick C, Mireskandari A, et al. Development and analytical validation of an expanded mutation detection panel for next-generation sequencing of thyroid nodule aspirates. *J Mol Diagn* 2020;22(3):355-67.
- Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med* 2012;367(8):705-15.
- Ali SZ, Siperstein A, Sadow PM, Golding AC, Kennedy GC, Kloos RT, Ladenson PW. Extending expressed RNA genomics from surgical decision making for cytologically indeterminate thyroid nodules to targeting therapies for metastatic thyroid cancer. *Cancer Cytopathol* 2019;127(6):362-9.
- Angell TE, Heller HT, Cibas E, Barletta J, Kim MI, Krane JF, Marqusee E. An independent comparison of the Afirma Genomic Sequencing Classifier (GSC) and Gene Expression Classifier (GEC) for cytologically indeterminate thyroid nodules. *Thyroid* 2019;29(5):650-6.
- Angell TE, Frates MC, Medici M, Liu X, Kwong N, Cibas ES, et al. Afirma benign thyroid nodules show similar growth to cytologically benign nodules during follow-up. *J Clin Endocrinol Metab* 2015;100(11):E1477-83.
- Arosemena M, Thekkumkattil A, Valderrama ML, Kuker R, Castillo RP, Sidani C, et al. American Thyroid Association sonographic risk and Afirma gene expression classifier alone and in combination for the diagnosis of thyroid nodules with Bethesda category III cytology. *Thyroid* 2020;30(11):1613-9.
- Balentine CJ, Vanness DJ, Schneider DF. Cost-effectiveness of lobectomy versus genetic testing (Afirma®) for indeterminate thyroid nodules: Considering the costs of surveillance. *Surgery* 2018;163(1):88-96.
- Beaudenon-Huibregtse S, Alexander EK, Guttler RB, Hershman JM, Babu V, Blevins TC, et al. Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. *Thyroid* 2014;24(10):1479-87.
- Bernet V, Hupart KH, Parangi S, Woeber KA. AACE/ACE disease state commentary: Molecular diagnostic testing of thyroid nodules with indeterminate cytopathology. *Endocr Pract* 2014;20(4):360-3.
- Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: A meta-analysis. *Acta Cytol* 2012;56(4):333-9.
- Chaudhary S, Hou Y, Shen R, Hooda S, Li Z. Impact of the Afirma gene expression classifier result on the surgical management of thyroid nodules with category III/IV cytology and its correlation with surgical outcome. *Acta Cytol* 2016;60(3):205-10.

- Chen T, Gilfix BM, Rivera J, Sadeghi N, Richardson K, Hier MP, et al. The role of the ThyroSeq v3 molecular test in the surgical management of thyroid nodules in the Canadian public health care setting. *Thyroid* 2020;30(9):1280-7.
- Cherella CE, Angell TE, Richman DM, Frates MC, Benson CB, Moore FD, et al. Differences in thyroid nodule cytology and malignancy risk between children and adults. *Thyroid* 2019;29(8):1097-104.
- Chudova D, Wilde JI, Wang ET, Wang H, Rabbee N, Egidio CM, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab* 2010;95(12):5296-304.
- Cibas ES et Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* 2017;27(11):1341-6.
- Cibas ES et Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* 2009;19(11):1159-65.
- Crochard-Lacour A et LeLorier J. Introduction à la pharmacoeconomie. Montréal, Qc : Presses de l'Université de Montréal; 2000.
- Deaver KE, Haugen BR, Pozdeyev N, Marshall CB. Outcomes of Bethesda categories III and IV thyroid nodules over 5 years and performance of the Afirma Gene Expression Classifier: A single-institution study. *Clin Endocrinol (Oxf)* 2018;89(2):226-32.
- Duh QY, Busaidy NL, Rahilly-Tierney C, Gharib H, Randolph G. A systematic review of the methods of diagnostic accuracy studies of the Afirma Gene Expression Classifier. *Thyroid* 2017;27(10):1215-22.
- Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: A review. *JAMA* 2018;319(9):914-24.
- Endo M, Nabhan F, Porter K, Roll K, Shirley LA, Azaryan I, et al. Afirma Gene Sequencing Classifier compared with gene expression classifier in indeterminate thyroid nodules. *Thyroid* 2019;29(8):1115-24.
- Ferris RL, Baloch Z, Bernet V, Chen A, Fahey TJ 3rd, Ganly I, et al. American Thyroid Association statement on surgical application of molecular profiling for thyroid nodules: Current impact on perioperative decision making. *Thyroid* 2015;25(7):760-8.
- Francis GL, Waguespack SG, Bauer AJ, Angelos P, Benvenga S, Cerutti JM, et al. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2015;25(7):716-59.
- Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedus L, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules – 2016 Update. *Endocr Pract* 2016;22(5):622-39.

- Giordano TJ, Beaudenon-Huibregtse S, Shinde R, Langfield L, Vinco M, Laosinchai-Wolf W, Labourier E. Molecular testing for oncogenic gene mutations in thyroid lesions: A case-control validation study in 413 postsurgical specimens. *Hum Pathol* 2014;45(7):1339-47.
- Guertin JR, Feeny D, Tarride JE. Age- and sex-specific Canadian utility norms, based on the 2013-2014 Canadian Community Health Survey. *CMAJ* 2018;190(6):E155-61.
- Hao Y, Choi Y, Babiarz JE, Kloos RT, Kennedy GC, Huang J, Walsh PS. Analytical verification performance of Afirma Genomic Sequencing Classifier in the diagnosis of cytologically indeterminate thyroid nodules. *Front Endocrinol (Lausanne)* 2019;10:438.
- Harrell RM, Eyerly-Webb SA, Golding AC, Edwards CM, Bimston DN. Statistical comparison of Afirma GSC and Afirma GEC outcomes in a community endocrine surgical practice: Early findings. *Endocr Pract* 2019;25(2):161-4.
- Haugen BR, Sawka AM, Alexander EK, Bible KC, Caturegli P, Doherty GM, et al. American Thyroid Association Guidelines on the Management of Thyroid Nodules and Differentiated Thyroid Cancer Task Force review and recommendation on the proposed renaming of encapsulated follicular variant papillary thyroid carcinoma without invasion to noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Thyroid* 2017;27(4):481-3.
- Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* 2016;26(1):1-133.
- Institut national d'excellence en santé et en services sociaux (INESSS). Glossaire en ETS [site Web]. Québec, Qc : INESSS; 2018. Disponible à : <http://htaglossary.net/Accueil>.
- Kandil E, Krishnan B, Noureldine SI, Yao L, Tufano RP. Hemithyroidectomy: A meta-analysis of postoperative need for hormone replacement and complications. *ORL J Otorhinolaryngol Relat Spec* 2013;75(1):6-17.
- Kay-Rivest E, Tibbo J, Bouhabel S, Tamilia M, Leboeuf R, Forest VI, et al. The first Canadian experience with the Afirma® gene expression classifier test. *J Otolaryngol Head Neck Surg* 2017;46(1):25.
- Kloos RT. Molecular profiling of thyroid nodules: Current role for the Afirma gene expression classifier on clinical decision making. *Mol Imaging Radionucl Ther* 2017;26(Suppl 1):36-49.
- Labourier E, Shifrin A, Busseniers AE, Lupo MA, Manganelli ML, Andruss B, et al. Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *J Clin Endocrinol Metab* 2015;100(7):2743-50.

- Lee DJ, Xu JJ, Brown DH, Gilbert RW, Gullane PJ, Irish JC, et al. Determining patient preferences for indeterminate thyroid nodules: Observation, surgery or molecular tests. *World J Surg* 2017;41(6):1513-20.
- Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab* 2011;96(11):E1719-26.
- Lithwick-Yanai G, Dromi N, Shtabsky A, Morgenstern S, Strenov Y, Feinmesser M, et al. Multicentre validation of a microRNA-based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate smears. *J Clin Pathol* 2017;70(6):500-7.
- Livhits MJ, Kuo EJ, Leung AM, Rao J, Levin M, Douek ML, et al. Gene expression classifier vs targeted next-generation sequencing in the management of indeterminate thyroid nodules. *J Clin Endocrinol Metab* 2018;103(6):2261-8.
- Marti JL, Avadhani V, Donatelli LA, Niyogi S, Wang B, Wong RJ, et al. Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules. *Ann Surg Oncol* 2015;22(12):3996-4001.
- Mclver B, Castro MR, Morris JC, Bernet V, Smallridge R, Henry M, et al. An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab* 2014;99(11):4069-77.
- Meltzer CJ, Irish J, Angelos P, Busaidy NL, Davies L, Dwojak S, et al. American Head and Neck Society Endocrine Section clinical consensus statement: North American quality statements and evidence-based multidisciplinary workflow algorithms for the evaluation and management of thyroid nodules. *Head Neck* 2019;41(4):843-56.
- Mitchell AL, Gandhi A, Scott-Coombes D, Perros P. Management of thyroid cancer: United Kingdom National Multidisciplinary Guidelines. *J Laryngol Otol* 2016;130(S2):S150-60.
- Mitchell J et Yip L. Decision making in indeterminate thyroid nodules and the role of molecular testing. *Surg Clin North Am* 2019;99(4):587-98.
- National Comprehensive Cancer Network (NCCN). Thyroid carcinoma. Version 1.2019. Clinical Practice Guidelines in Oncology. Plymouth Meeting, PA : NCCN; 2019. Disponible à : https://www.nccn.org/professionals/physician_gls/.
- Nicholson K, Roberts MS, McCoy KL, Carty SE, Yip L. Molecular testing versus diagnostic lobectomy in Bethesda III/IV thyroid nodules: A cost-effectiveness analysis. *Thyroid* 2019;29(9):1237-43.
- Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, et al. Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid* 2015;25(11):1217-23.

- Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer* 2014;120(23):3627-34.
- Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: A prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab* 2011;96(11):3390-7.
- Nikiforov YE, Steward DL, Robinson-Smith TM, Haugen BR, Kloppner JP, Zhu Z, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab* 2009;94(6):2092-8.
- Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer* 2018;124(8):1682-90.
- Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab* 2013;98(11):E1852-60.
- Nishino M et Nikiforova M. Update on molecular testing for cytologically indeterminate thyroid nodules. *Arch Pathol Lab Med* 2018;142(4):446-57.
- Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, et al. SEER Cancer Statistics Review, 1975-2015. Bethesda, MD : National Cancer Institute; 2018. Disponible à : https://seer.cancer.gov/csr/1975_2015/.
- Nouraei SA, Virk JS, Middleton SE, Aylin P, Mace A, Vaz F, et al. A national analysis of trends, outcomes and volume-outcome relationships in thyroid surgery. *Clin Otolaryngol* 2017;42(2):354-65.
- Ohori NP, Landau MS, Carty SE, Yip L, LeBeau SO, Manroa P, et al. Benign call rate and molecular test result distribution of ThyroSeq v3. *Cancer Cytopathol* 2019;127(3):161-8.
- Padmanabhan V, Marshall CB, Akdas Barkan G, Ghofrani M, Laser A, Tolgay Ocal I, et al. Reproducibility of atypia of undetermined significance/follicular lesion of undetermined significance category using the Bethesda System for Reporting Thyroid Cytology when reviewing slides from different institutions: A study of interobserver variability among cytopathologists. *Diagn Cytopathol* 2017;45(5):399-405.
- Paschke R, Cantara S, Crescenzi A, Jarzab B, Musholt TJ, Sobrinho Simoes M. European Thyroid Association guidelines regarding thyroid nodule molecular fine-needle aspiration cytology diagnostics. *Eur Thyroid J* 2017;6(3):115-29.
- Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg* 2018;153(9):817-24.

- Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard G, et al. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)* 2014;81(Suppl 1):1-122.
- Sacks WL, Bose S, Zumsteg ZS, Wong R, Shiao SL, Braunstein GD, Ho AS. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. *Cancer Cytopathol* 2016;124(10):722-8.
- Sahli ZT, Smith PW, Umbricht CB, Zeiger MA. Preoperative molecular markers in thyroid nodules. *Front Endocrinol (Lausanne)* 2018;9:179.
- Samulski TD, LiVolsi VA, Wong LQ, Baloch Z. Usage trends and performance characteristics of a "gene expression classifier" in the management of thyroid nodules: An institutional experience. *Diagn Cytopathol* 2016;44(11):867-73.
- San Martin VT, Lawrence L, Bena J, Madhun NZ, Berber E, Elsheikh TM, Nasr CE. Real world comparison of Afirma GEC and GSC for the assessment of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab* 2020;105(3):e428-35.
- Singer J, Hanna JW, Visaria J, Gu T, McCoy M, Kloos RT. Impact of a gene expression classifier on the long-term management of patients with cytologically indeterminate thyroid nodules. *Curr Med Res Opin* 2016;32(7):1225-32.
- Sipos JA, Blevins TC, Shea HC, Duick DS, Lakhian SK, Michael BE, et al. Long-term nonoperative rate of thyroid nodules with benign results on the Afirma gene expression classifier. *Endocr Pract* 2016;22(6):666-72.
- Société canadienne du cancer (SCC). Statistiques canadiennes sur le cancer 2019. Toronto, ON : SCC; 2019. Disponible à : <https://www.cancer.ca/~media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2019-FR.pdf>.
- Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sipos JA, et al. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: A prospective blinded multicenter study. *JAMA Oncol* 2019;5(2):204-12.
- Sullivan PS, Hirschowitz SL, Fung PC, Apple SK. The impact of atypia/follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda System. *Cancer Cytopathol* 2014;122(12):866-72.
- Sultan R, Levy S, Sulanc E, Honasoge M, Rao SD. Utility of Afirma gene expression classifier for evaluation of indeterminate thyroid nodules and correlation with ultrasound risk assessment: Single institutional experience. *Endocr Pract* 2020;26(5):543-51.
- Taye A, Gurciullo D, Miles BA, Gupta A, Owen RP, Inabnet WB 3rd, et al. Clinical performance of a next-generation sequencing assay (ThyroSeq v2) in the evaluation of indeterminate thyroid nodules. *Surgery* 2018;163(1):97-103.

- Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J Am Coll Radiol* 2017;14(5):587-95.
- Topstad D et Dickinson JA. Thyroid cancer incidence in Canada: A national cancer registry analysis. *CMAJ Open* 2017;5(3):E612-6.
- Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L. Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. *N Engl J Med* 2016;375(7):614-7.
- Valderrabano P, Hallanger-Johnson JE, Thapa R, Wang X, McIver B. Comparison of postmarketing findings vs the initial clinical validation findings of a thyroid nodule gene expression classifier: A systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg* 2019;145(9):783-92.
- Valderrabano P, Khazai L, Leon ME, Thompson ZJ, Ma Z, Chung CH, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. *Endocr Relat Cancer* 2017;24(3):127-36.
- Verloop H, Louwerens M, Schoones JW, Kievit J, Smit JW, Dekkers OM. Risk of hypothyroidism following hemithyroidectomy: Systematic review and meta-analysis of prognostic studies. *J Clin Endocrinol Metab* 2012;97(7):2243-55.
- Villabona CV, Mohan V, Arce KM, Diacovo J, Aggarwal A, Betancourt J, et al. Utility of ultrasound versus gene expression classifier in thyroid nodules with atypia of undetermined significance. *Endocr Pract* 2016;22(10):1199-203.
- Wei S, Veloski C, Sharda P, Ehya H. Performance of the Afirma genomic sequencing classifier versus gene expression classifier: An institutional experience. *Cancer Cytopathol* 2019;127(11):720-4.
- Weiss A, Parina RP, Tang JA, Brumund KT, Chang DC, Bouvet M. Outcomes of thyroidectomy from a large California state database. *Am J Surg* 2015;210(6):1170-7.
- Wong CW, Schumm MA, Zhu CY, Tseng CH, Arasu A, Han J, Livhits MJ. Quality of life following molecular marker testing for indeterminate thyroid nodules. *Endocr Pract* 2020 [Epub ahead of print].
- Wu JX, Young S, Hung ML, Li N, Yang SE, Cheung DS, et al. Clinical factors influencing the performance of gene expression classifier testing in indeterminate thyroid nodules. *Thyroid* 2016;26(7):916-22.
- Yang SE, Sullivan PS, Zhang J, Govind R, Levin MR, Rao JY, Moatamed NA. Has Afirma gene expression classifier testing refined the indeterminate thyroid category in cytology? *Cancer Cytopathol* 2016;124(2):100-9.
- Zanocco KA, Wang MM, Yeh MW, Livhits MJ. Selective use of molecular testing based on sonographic features of cytologically indeterminate thyroid nodules: A decision analysis. *World J Surg* 2020;44(2):393-401.

ANNEXE A

Stratégie de repérage de l'information scientifique

Date de la recherche : 19 mars 2019

Limites : 2009 -; anglais, français

PubMed (NLM)	
#1	thyroid[tiab] AND nodule*[tiab] AND (indeterminate[tiab] OR bethesda[tiab]) AND (molecular[tiab] OR afirma[tiab] OR classifier[tiab] OR thyroseq[tiab] OR thygenx[tiab] OR thyramir[tiab] OR rosetta[tiab])
#2	"Thyroid Nodule/analysis"[majr] OR "Thyroid Nodule/cytology"[majr] OR "Thyroid Nodule/diagnosis"[majr]
#3	accurac*[tiab] OR (clinical[tiab] AND trial[tiab]) OR compar*[tiab] OR correlation[tiab] OR correlated[tiab] OR (criter*[tiab] AND (bias[tiab] OR standard*[tiab] OR test[tiab] OR testing[tiab] OR tests[tiab])) OR curve[tiab] OR curves[tiab] OR decision making[tiab] OR detect*[tiab] OR diagnos*[tiab] OR diagnostic[tiab] OR differential*[tiab] OR distinguish*[tiab] OR enhancement[tiab] OR expectation bias[tiab] OR identif*[tiab] OR indeterminate result*[tiab] OR (index[tiab] AND (test[tiab] OR testing[tiab] OR tests[tiab])) OR interater reliability[tiab] OR interrater reliability[tiab] OR likelihood ratio*[tiab] OR logistic model*[tiab] OR logistic regression[tiab] OR logistical regression[tiab] OR false negative*[tiab] OR false positive*[tiab] OR (false[tiab] AND reaction*[tiab]) OR gold standard*[tiab] OR misdiagnos*[tiab] OR monitoring[tiab] OR negative rate*[tiab] OR (observer[tiab] AND (bias[tiab] OR variat*[tiab])) OR observer reliability[tiab] OR (performance[tiab] AND (aid[tiab] OR aids[tiab] OR score*[tiab] OR standard*[tiab] OR test[tiab] OR testing[tiab] OR tests[tiab] OR tool*[tiab])) OR predict*[tiab] OR positive rate*[tiab] OR ((post-test[tiab] OR posttest[tiab] OR pre-test[tiab] OR pretest[tiab]) AND probabiliti*[tiab]) OR random*[tiab] OR (receiver operat*[tiab] AND (characteristic*[tiab] OR curve[tiab])) OR (reference[tiab] AND (test[tiab] OR testing[tiab] OR tests[tiab])) OR reference standard*[tiab] OR reference value*[tiab] OR ((reliability[tiab] OR reliable[tiab]) AND (aid[tiab] OR aids[tiab] OR intra[tiab] OR score*[tiab] OR standard*[tiab] OR test[tiab] OR testing[tiab] OR tests[tiab] OR tool*[tiab])) OR reproducibility[tiab] OR reproductivity[tiab] OR roc-curve[tiab] OR "roc"[tiab] OR "rocs"[tiab] OR "sroc"[tiab] OR "srocs"[tiab] OR screen*[tiab] OR sensitiv*[tiab] OR specificity[tiab] OR ((single[tiab] OR double[tiab] OR triple[tiab]) AND blind*[tiab]) OR test outcome[tiab] OR true negative*[tiab] OR true positive*[tiab] OR (validat*[tiab] AND (bias[tiab] OR standard*[tiab] OR test[tiab] OR testing[tiab] OR tests[tiab])) OR (verificat* AND bias[tiab])
#4	(indeterminate[tiab] OR bethesda[tiab]) AND (molecular[tiab] OR afirma[tiab] OR classifier[tiab] OR thyroseq[tiab] OR thygenx[tiab] OR thyramir[tiab] OR rosetta[tiab])
#5	#1 OR (#2 AND #3 AND #4)

* troncature; / qualificatif de descripteur; [mh] descripteur sujet; [mh:noexp] descripteur sujet sans inclusion de ses termes spécifiques; [pt] type de publication; [sh] qualificatif de descripteur; [ti] titre; [tiab] titre ou résumé; .tw combinaison de champs textes.

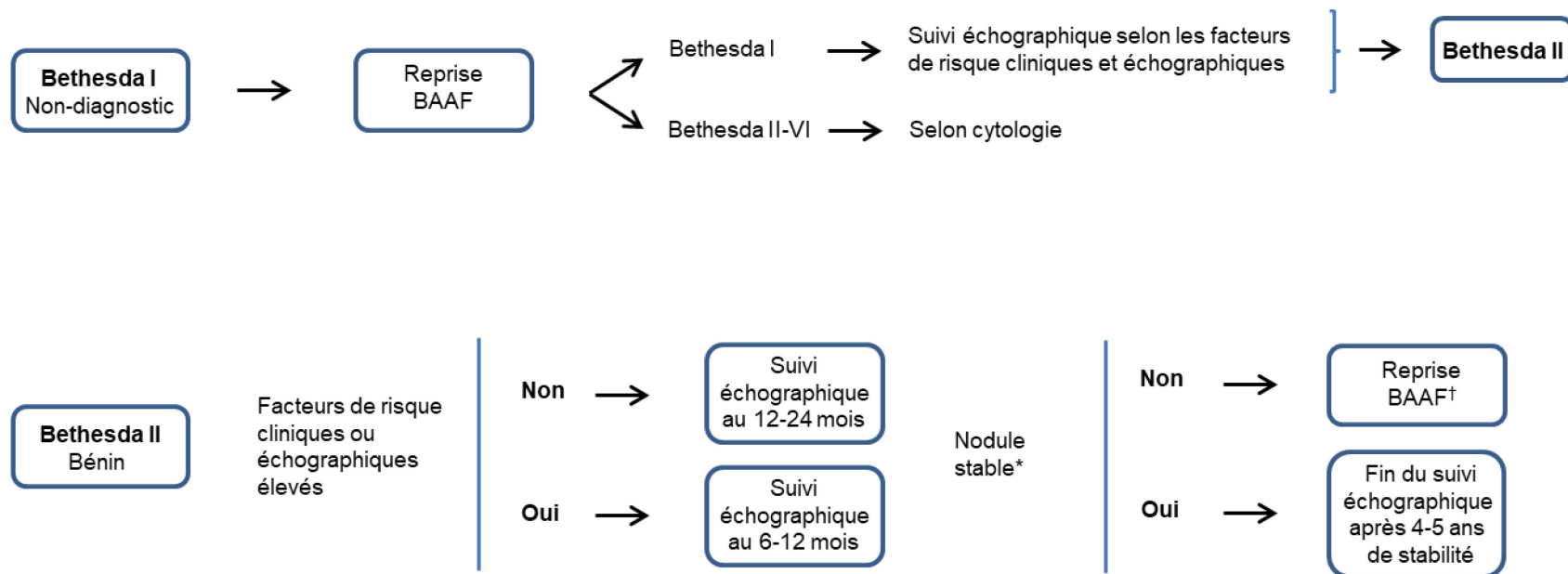
Embase (Ovid)	
1	*thyroid nodule/di [Diagnosis]
2	(thyroid* AND nodule*).ti
3	(afirma OR thyroseq OR thygenx OR rosetta OR molecular OR classifier).tw
4	(indeterminate OR bethesda).ti
5	(1 OR 2) AND (3 AND 4)
EBM Reviews (Ovid)	
Cochrane Database of Systematic Reviews; Health Technology Assessment; NHS Economic Evaluation Database	
1	thyroid*.ti. AND nodule*.ti.ab. AND indetermin*.tw

* troncature à la fin d'un mot; * sujet principal devant un descripteur; / descripteur; /di qualificatif de descripteur Diagnosis; .ab résumé; .tw combinaison de champs texte.

ANNEXE B

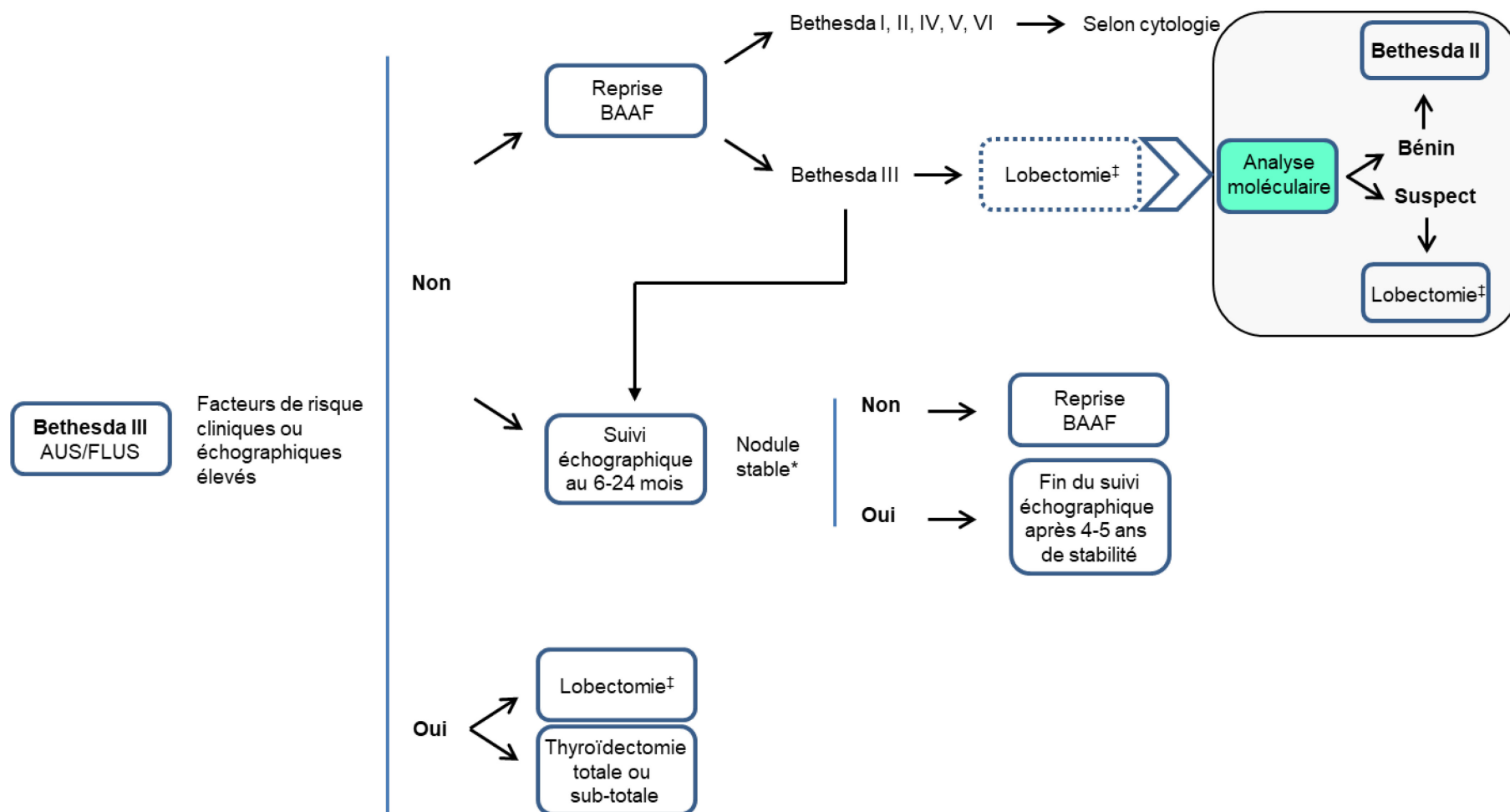
Algorithme décisionnel

Figure B-1 Prise en charge des nodules thyroïdiens selon le résultat de la cytologie



* En l'absence d'une croissance significative définie comme une augmentation d'au moins 20 % dans un minimum de 2 dimensions ou d'une augmentation du volume d'au moins 50 %.

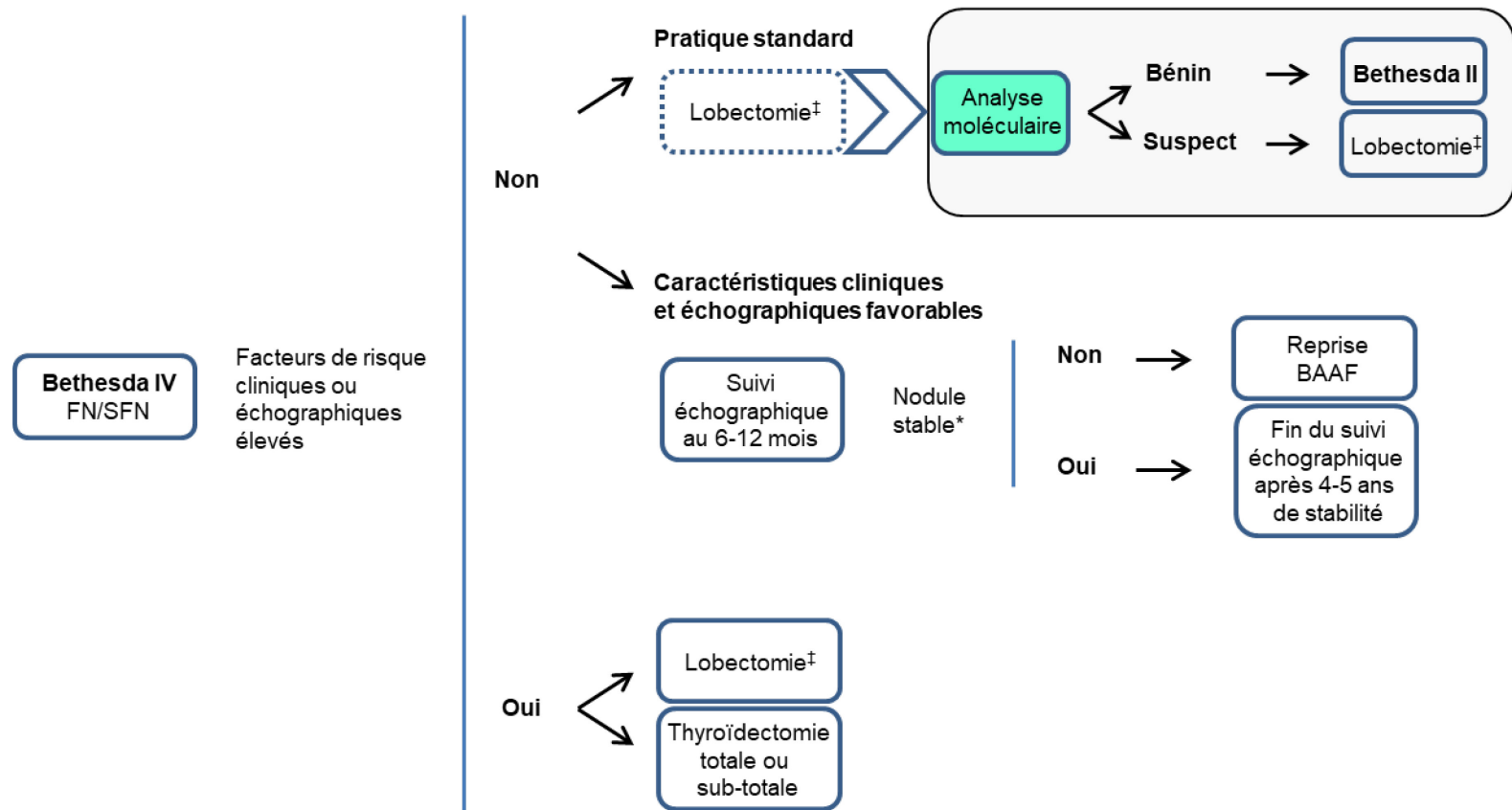
† Si un nodule soumis à une reprise de la BAAF présente un second résultat de cytologie bénigne, la surveillance échographique de ce nodule n'est plus indiquée.



AUS/FLUS : atypie de signification indéterminée ou lésion folliculaire de signification indéterminée.

* En l'absence d'une croissance significative définie comme une augmentation d'au moins 20 % dans un minimum de 2 dimensions ou d'une augmentation du volume d'au moins 50 %.

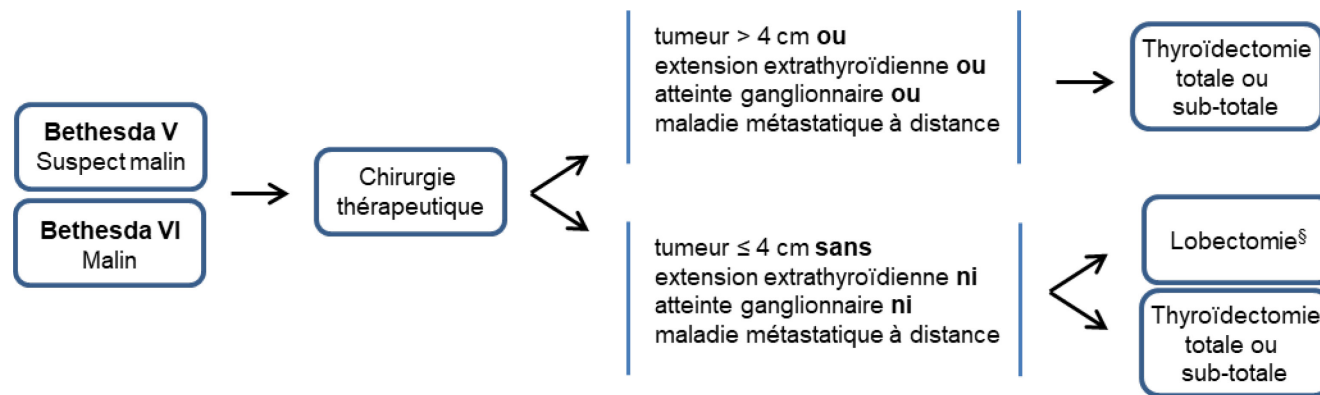
‡ Une reprise de la thyroïdectomie pour goitre restant (complétion) est pratiquée lorsqu'un carcinome papillaire est décelé suivant la lobectomie et qu'un des éléments suivants est identifié: la tumeur est > 4 cm, les marges de résection sont positives, en présence d'une extension extrathyroïdienne macroscopique, d'une atteinte ganglionnaire, d'une maladie controlatérale, d'une invasion vasculaire ou d'un cancer peu différencié.



FN/SFN : néoplasme folliculaire ou suspicion d'un néoplasme folliculaire.

* En l'absence d'une croissance significative définie comme une augmentation d'au moins 20 % dans un minimum de 2 dimensions ou d'une augmentation du volume d'au moins 50 %.

‡ Une reprise de la thyroïdectomie pour goitre restant (complétion) est pratiquée lorsqu'un carcinome papillaire est décelé suivant la lobectomie et qu'un des éléments suivants est identifié: la tumeur est > 4 cm, les marges de résection sont positives, en présence d'une extension extrathyroïdienne macroscopique, d'une atteinte ganglionnaire, d'une maladie controlatérale, d'une invasion vasculaire ou d'un cancer peu différencié.



§ Une reprise de la thyroïdectomie pour goitre restant (complétion) est pratiquée après la lobectomie lorsqu'un des éléments suivants est identifié: la tumeur est > 4 cm, les marges de résection sont positives, en présence d'une extension extrathyroïdienne macroscopique, d'une atteinte ganglionnaire, d'une maladie controlatérale, d'une invasion vasculaire ou d'un cancer peu différencié.

ANNEXE C

Sélection des études

Figure C-1 Diagramme de flux

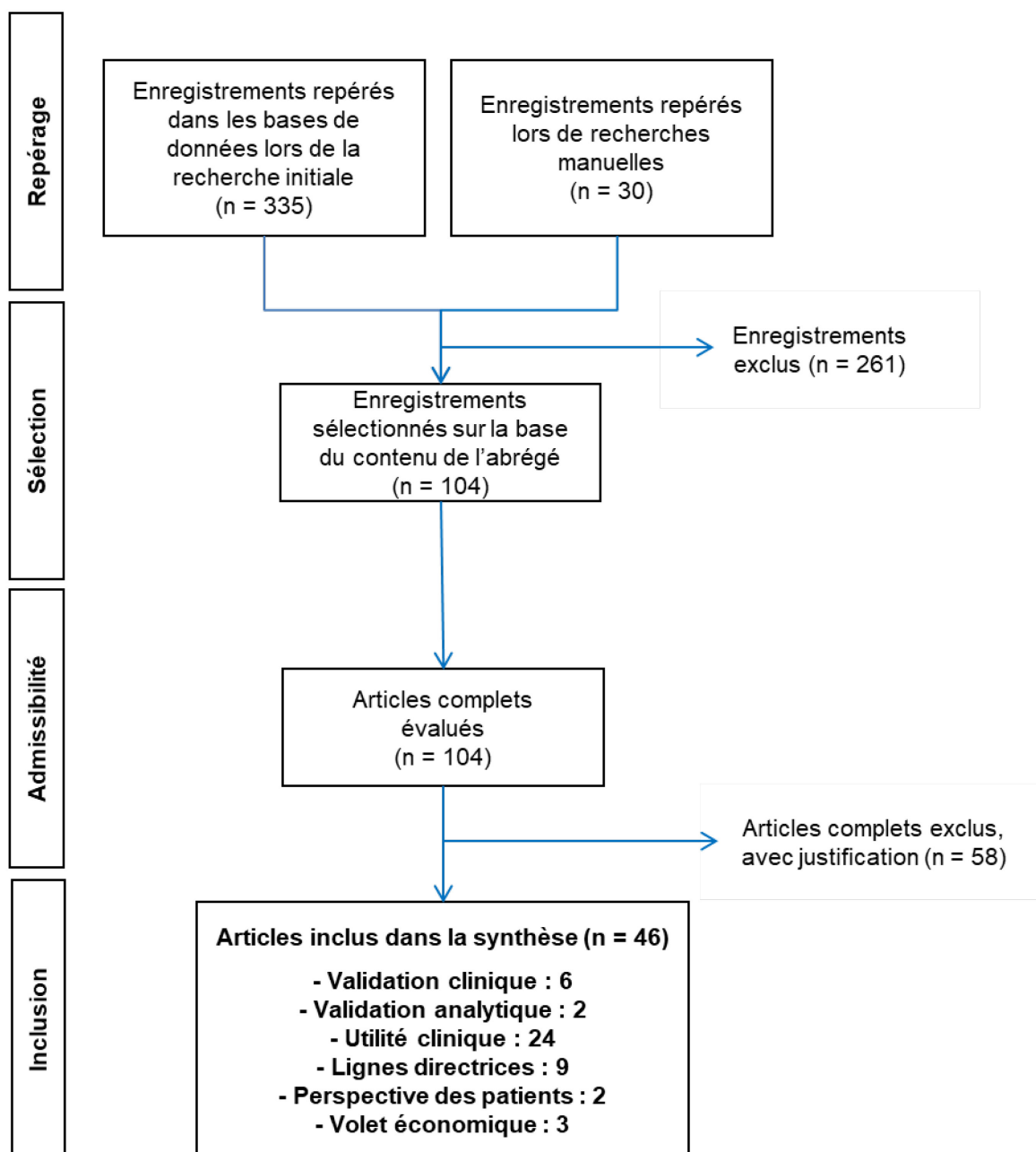


Tableau C-1 Liste des études exclues et principales raisons de l'exclusion

ID	Études et méta-analyses	Principales raisons de l'exclusion
1	Vora <i>et al.</i> , Otolaryngol Head Neck Surg 2020 (sélection manuelle)	Les données sont présentées pour l'ensemble des nodules à cytologie des catégories I-V (n = 416); le résultat de la cytologie n'était pas indéterminé pour 4 à 5 % des nodules.
2	Sistrunk <i>et al.</i> , J Am Soc Cytopathol 2020 (sélection manuelle)	Une étude de validation clinique comprenant une évaluation de la sensibilité, de la spécificité et des valeurs prédictives positive et négative de l'analyse ThyGeNext/ThyraMIR n'a pas encore été publiée. L'étude menée par Sistrunk comprend une évaluation de la performance clinique basée sur la probabilité de survie sans maladie/malignité selon le résultat du test (analyse Kaplan Meier).
3	Jug <i>et al.</i> , Cancer Cytopathol 2020 (sélection manuelle)	Les nodules pédiatriques étaient inclus (9-87 ans).
4	Fazeli <i>et al.</i> , Thyroid 2020 (sélection manuelle)	Le modèle économique est basé sur la comparaison de deux cohortes de patients évalués avant ou après l'introduction de l'analyse ThyroSeq v2; l'analyse inclut les nodules des catégories Bethesda I-VI. Parmi les nodules à cytologie indéterminée de la période post-ThyroSeq v2 (n = 105), 36 % se sont avérés à cytologie bénigne après une 2 ^e évaluation cytopathologique. La prise en charge subséquente n'a toutefois pas été ajustée en fonction du résultat de la 2 ^e évaluation. Les résultats des analyses moléculaires n'ont pas été présentés.
5	Andrioli <i>et al.</i> , Endocr Pathol 2020 (sélection manuelle)	Les auteurs utilisent un système de classification des résultats de cytologie (ICCRTC : <i>Italian consensus for the classification and reporting of thyroid cytology</i>) différent du système de classification Bethesda.
6	Sistrunk <i>et al.</i> , Diagn Cytopathol 2019 (sélection manuelle)	Les auteurs indiquent avoir évalué l'utilité clinique de l'analyse ThyGenX/ThyraMIR telle que décrite par Labourier et ses collaborateurs en 2015. Cependant, la version utilisée est vraisemblablement différente, puisqu'elle comprend le séquençage du gène <i>PIK3CA</i> , non réalisé dans la version originale. Le suivi des nodules après l'évaluation moléculaire n'était pas documenté pour 47 % des patients, exclus de l'analyse.
7	Valderrabano <i>et al.</i> , JAMA Otolaryngol Head Neck Surg 2019 (sélection manuelle)	Méta-analyse réalisée dans le but d'évaluer la performance du test Afirma GEC et comportant des limitations en lien avec la sélection des études incluses dans la méta-régression (p. ex. nodules sélectionnés pour une résection chirurgicale et inclusion d'un certain nombre de nodules pédiatriques).
8	Partyka <i>et al.</i> , Diagn Cytopathol 2019 (sélection manuelle)	Le nombre de nodules examinés était limité* (Afirma GEC/GSC combinés : n = 68; ThyGenX/ThyraMIR : n = 22). Certains nodules étaient pédiatriques; la cytologie n'était pas indéterminée pour tous les nodules.
9	Marcadis <i>et al.</i> , Surgery 2019	L'évaluation de la performance de ThyroSeq v2 est basée sur des nodules à cytologie indéterminée sélectionnés pour une résection chirurgicale.
10	Liu <i>et al.</i> , Biomed Res Int 2019 (sélection manuelle)	Méta-analyse réalisée dans le but d'évaluer la performance du test Afirma GEC et comportant des limitations en lien avec la sélection des études incluses dans la méta-régression (p. ex. nodules sélectionnés pour une résection chirurgicale, inclusion d'un certain nombre de nodules dont le résultat de la cytologie n'était pas indéterminé, certains nodules étaient pédiatriques).
11	Jug <i>et al.</i> , Endocr Pathol 2019	Les nodules pédiatriques étaient inclus (âges non spécifiés). Les résultats pour le test ThyroSeq combinent ceux des versions v0, v1 et v2.
12	Ohori <i>et al.</i> , Cancer Cytopathol 2019 (sélection manuelle)	Les données sur la résection chirurgicale ne sont pas présentées en fonction du résultat au test moléculaire.
13	Young <i>et al.</i> , J Am Soc Cytopathol 2018	Les données sont présentées pour l'ensemble des nodules à cytologie des catégories I à VI (n = 95); le résultat de la cytologie n'était pas indéterminé pour 8 nodules. Trois tests moléculaires ont été utilisés : Afirma GEC (n = 17), Thyroseq v2 (n = 54) et ThyGenX/ThyraMIR (n = 24).
14	Walts <i>et al.</i> , Diagn Cytopathol 2018	L'évaluation de la performance de RosettaGX Reveal et Afirma GEC est basée sur des nodules sélectionnés pour une résection chirurgicale. Les nodules pédiatriques étaient inclus (14-81 ans).

ID	Études et méta-analyses	Principales raisons de l'exclusion
15	Vargas-Salas <i>et al.</i> , Endocr Relat Cancer 2018	Méta-analyse réalisée dans le but d'évaluer la performance des tests ThyroSeq v2 et Afirma GEC et comportant des limitations en lien avec la sélection des études incluses dans la méta-régression (p. ex. nodules sélectionnés pour une résection chirurgicale et inclusion d'un certain nombre de nodules dont le résultat de la cytologie n'était pas indéterminé).
16	Santos <i>et al.</i> , Thyroid 2018	Publication unique sur le développement et l'évaluation de la performance d'un test moléculaire (mir-THYpe) offert par une entreprise privée établie au Brésil.
17	Rivas <i>et al.</i> , Endocr Pract 2018	L'évaluation économique est basée sur les données d'un nombre limité de nodules : 8 nodules à cytologie Bethesda III et 13 nodules à cytologie Bethesda IV soumis à l'évaluation moléculaire avec ThyroSeq v2 ont été comparés à 8 nodules à cytologie Bethesda III et 11 nodules à cytologie Bethesda IV soumis à une résection chirurgicale. De plus, les patients du groupe BIV sans évaluation moléculaire répondaient à une indication de chirurgie ou avaient une préférence pour cette approche.
18	Jug <i>et al.</i> , Cancer Cytopathol 2018	Les résultats pour le test ThyroSeq combinent ceux des versions v0 et v2.
19	Harrell <i>et al.</i> , Endocr Pract 2018	Les nodules de ces patients ont aussi été évalués par Harrell et ses collaborateurs [2019] dont les résultats sont rapportés dans le présent document.
20	Borrowczyk <i>et al.</i> , Endocr Pathol 2018	Méta-analyse réalisée dans le but d'évaluer la performance des tests ThyroSeq v2 et Afirma GEC et comportant des limitations en lien avec la sélection des études incluses dans la méta-régression (p. ex. nodules sélectionnés pour une résection chirurgicale, inclusion d'un certain nombre de nodules dont le résultat de la cytologie n'était pas indéterminé, certains nodules étaient pédiatriques).
21	Shapiro <i>et al.</i> , Otolaryngol Head Neck Surg 2017	Le modèle économique est basé sur les données de suivi d'une cohorte de patients évalués entre 2009 et 2014 selon une version antérieure des recommandations de l'ATA (2009), qui diffère de celle présentement en vigueur (2015) quant à la prise en charge des nodules malins.
22	Kloos, Mol Imaging Radionucl Ther 2017	Revue narrative
23	Sciaccitano <i>et al.</i> , Oncotarget 2017	Dans cette publication, les tests moléculaires d'intérêt pour la présente évaluation n'ont pas fait l'objet d'une méta-analyse; seules les données des études de validation clinique originales sont présentées.
24	Harrison <i>et al.</i> , Arch Pathol Lab Med 2017	Les nodules pédiatriques étaient inclus (16-87 ans).
25	Hang <i>et al.</i> , Cancer Cytopathol 2017	Les nodules pédiatriques étaient inclus (8-90 ans).
26	Duh <i>et al.</i> , Thyroid 2017	Aucune donnée quantitative; l'article porte sur l'évaluation de la qualité méthodologique des études qui ont évalué la performance du test GEC (Afirma).
27	Baca <i>et al.</i> , Cancer Cytopathol 2017	Les nodules de ces patients ont aussi été évalués par Angell et ses collaborateurs [2019] dont les résultats sont rapportés dans le présent document.
28	Al-Qurayshi <i>et al.</i> , JAMA Otolaryngol Head Neck Surg 2017	L'utilité clinique du test moléculaire ne peut être évaluée dans le contexte où les données suggèrent l'inclusion de nodules pour lesquels une chirurgie était initialement indiquée : les auteurs précisent que tous les patients au nodule GEC bénin répondaient à l'indication d'une chirurgie (dont 62 % ont été réséqués); le délai médian entre le test moléculaire et la chirurgie a été de 30 jours (EIQ 20-63) pour les nodules GEC bénins.
29	Wu <i>et al.</i> , Surgery 2016	Le modèle économique est basé sur les données de suivi d'une cohorte de patients évalués entre 2012 et 2014 selon une version antérieure des recommandations de l'ATA (2009), qui diffère de celle présentement en vigueur (2015) quant à la prise en charge des nodules malins.
30	Witt <i>et al.</i> , Del Med J 2016	Étude incluant peu de sujets (n = 26) et dont les résultats limités sur la perspective des patients (objectif principal) apportent peu à la présente évaluation.
31	Witt <i>et al.</i> , Laryngoscope 2016	Le nombre de nodules examinés est limité* (n = 32 dont 29 avec un résultat au test; 14 moléculairement bénins).
32	Valderrabano <i>et al.</i> , Eur J Endocrinol 2016	Évaluation de la performance du test miRInform (devenu ThyGenX); un <i>rule-in test</i> . Les données présentées pour Afirma et ThyroSeq v2 sont théoriques [†] .

ID	Études et méta-analyses	Principales raisons de l'exclusion
33	Shrestha <i>et al.</i> , Thyroid 2016	L'évaluation de la performance du test ThyroSeq est basée sur des nodules à cytologie indéterminée sélectionnés pour une résection chirurgicale (n = 56). L'étude comprend des nodules soumis à différentes versions du test (n = 17 pour v0 et n = 39 pour v1 ou v2).
34	Santhanam <i>et al.</i> , Med Oncol 2016	Méta-analyse réalisée dans le but d'évaluer la performance du test Afirma GEC et comportant des limitations en lien avec la sélection des études incluses dans la méta-régression (p. ex. nodules sélectionnés pour une résection chirurgicale et inclusion d'un certain nombre de nodules dont le résultat de la cytologie n'était pas indéterminé).
35	Reed <i>et al.</i> , Head Neck 2016	Évaluation théorique [†] de la valeur prédictive négative du test Afirma GEC. La prévalence de la maladie a été déterminée à partir d'une cohorte de patients sélectionnés pour la chirurgie.
36	Nourelidine <i>et al.</i> , JAMA Otolaryngol Head Neck Surg 2016	Comparaison de la prise en charge chirurgicale des patients qui présentent un nodule thyroïdien évalué à l'aide d'un test moléculaire (n = 140) avec ceux chez qui le test n'a pas été réalisé (n = 548). Sélection des patients orientés vers une consultation en chirurgie. L'étude comprend toutes les catégories cytologiques (I-VI), incluant celles qui ne sont pas indéterminées. Les nodules pédiatriques étaient inclus (9-93 ans). Le groupe de patients soumis au test moléculaire combine les données de tests divers à la fonction différente (<i>rule-in / rule-out</i>)
37	Labourier <i>et al.</i> , Clin Endocrinol 2016	Le modèle économique comprend la possibilité d'une thyroïdectomie totale en présence d'un résultat moléculairement suspect alors que la lobectomie est le seul type de chirurgie recommandé pour les nodules à cytologie indéterminée en l'absence de facteur de risque élevé. Le modèle incluait aussi une reprise de thyroïdectomie (chirurgie de complétion) pour tous les cas de cancer décelés après lobectomie, bien que les lignes directrices actuelles ne le recommandent que dans certains cas précis.
38	Jiang <i>et al.</i> , Cancer Cytopathol 2016	Données limitées*; 8 nodules à cytologie indéterminée (histopathologie NIFTP) dont 4 ont été analysés avec Afirma GEC et 4 avec ThyroSeq v2.
39	Dhingra <i>et al.</i> , Otolaryngol Head Neck Surg 2016	Étude comparant le taux de chirurgies diagnostiques pour les nodules à cytologie indéterminée avant et après l'introduction du test Afirma GEC. Le nombre de nodules examinés est limité* (n = 43 vs 27, dont 24 avec un résultat au test; 12 moléculairement bénins). La période d'introduction du test Afirma GEC est également marquée par la centralisation des BAAF, ce qui pourrait affecter l'interprétation des résultats.
40	Abeykoon <i>et al.</i> , Horm Cancer 2016	Étude comparant le taux de chirurgies diagnostiques pour les nodules à cytologie indéterminée avant et après l'introduction du test Afirma GEC. Le nombre de nodules examinés est limité* (n = 27 vs 34, dont 17 résultats moléculairement bénins). Les nodules pédiatriques étaient inclus (14-89 ans).
41	Zhu <i>et al.</i> , AJR Am J Roentgenol 2015	Le nombre de nodules examinés est limité* (n = 45 dont 44 avec un résultat au test; 23 moléculairement bénins).
42	Nourelidine <i>et al.</i> , JAMA Otolaryngol Head Neck Surg 2015	Étude évaluant l'impact du test Afirma GEC sur la prise en charge chirurgicale des nodules thyroïdiens. Sélection des patients orientés vers une consultation en chirurgie; sous-représentation des nodules dont le résultat au test est moléculairement bénin (n = 18 vs 197 résultats suspects, pour les nodules à cytologie indéterminée). Plus de 80 % des nodules ont été réséqués.
43	Celik <i>et al.</i> , Diagn Cytopathol 2015 (sélection manuelle)	Le nombre de nodules examinés est limité* (n = 66 dont 40 à cytologie indéterminée et seulement 34 avec un résultat au test; 11 moléculairement bénins).
44	Brauner <i>et al.</i> , Thyroid 2015	Évaluation limitée aux nodules à cytologie indéterminée de type oncocytaire (prédominance de cellules de Hürthle).
45	Lee <i>et al.</i> , J Clin Endocrinol Metab 2014 (sélection manuelle)	Le modèle économique est basé sur une version antérieure des recommandations de l'ATA (2009), qui diffère de celle présentement en vigueur (2015) quant à la prise en charge des nodules malins.
46	Lastra <i>et al.</i> , Cancer Cytopathol 2014	Les nodules de ces patients ont aussi été évalués par Samulski et ses collaborateurs [2016] dont les résultats sont rapportés dans le présent document.
47	Harrell <i>et al.</i> , Endocr Pract 2014	Les nodules de ces patients ont aussi été évalués par Harrell et ses collaborateurs [2019] dont les résultats sont rapportés dans le présent document.

ID	Études et méta-analyses	Principales raisons de l'exclusion
48	Giordano <i>et al.</i> , Hum Pathol 2014 (sélection manuelle)	Évaluation de la performance d'une version préliminaire (miRInform) de ThyGenX/ThyraMIR utilisée comme <i>rule-in test</i> .
49	Beaudenon-Huibregtse <i>et al.</i> , Thyroid 2014	Évaluation de la performance d'une version préliminaire (miRInform) de ThyGenX/ThyraMIR utilisée comme <i>rule-in test</i> .
50	Alexander <i>et al.</i> , J Clin Endocrinol Metab 2014	Les données sont présentées pour l'ensemble des nodules à cytologie de catégorie III à V (n = 339); le résultat de la cytologie n'était pas indéterminé pour 13 nodules. La présentation des données pour les nodules GEC bénins ne permettait pas l'exclusion des nodules à cytologie de catégorie V. Le suivi des nodules GEC bénins n'était pas documenté pour 59 % des patients.
51	Aragon Han <i>et al.</i> , Ann Surg Oncol 2014 (sélection manuelle)	Évaluation de l'impact des tests moléculaires sur la prise en charge chirurgicale des nodules thyroïdiens. Sélection des patients orientés vers une consultation en chirurgie dont le nodule a été réséqué (n = 87). Le résultat de la cytologie n'était pas indéterminé pour la majorité des patients (66 %). Les tests utilisés étaient divers et leur fonction différente (<i>rule-in / rule-out</i>).
52	Nikiforova <i>et al.</i> , J Clin Endocrinol Metab 2013 (sélection manuelle)	Évaluation de la performance d'une version préliminaire (v1) de ThyroSeq utilisée comme <i>rule-in test</i> .
53	Yip <i>et al.</i> , J Clin Endocrinol Metab 2012	Le modèle employé pour estimer l'impact économique des analyses moléculaires repose sur les données (coût/sensibilité/spécificité) d'une version préliminaire de ThyroSeq utilisée comme <i>rule-in test</i> .
54	Najafzadeh <i>et al.</i> , Value Health 2012	L'étude économique porte sur l'évaluation d'une analyse moléculaire hypothétique dont les caractéristiques de performance sont différentes de celles des tests évalués : sensibilité et spécificité de 95 %; valeur prédictive positive et négative de 92 et 97 %, respectivement. De plus, la catégorie indéterminée inclut les nodules soupçonnés de malignité (Bethesda V). Enfin l'analyse moléculaire n'est pas comparée à la lobectomie diagnostique mais plutôt à une approche globale combinant la biopsie à la classification selon le système Bethesda en l'absence de test moléculaire.
55	Duick <i>et al.</i> , Thyroid 2012	La catégorie « nodules à cytologie indéterminée » (n = 395) comprend un nombre inconnu de nodules soupçonnés de malignité (Bethesda V).
56	Nikiforov <i>et al.</i> , J Clin Endocrinol Metab 2011	Évaluation de la performance d'une version préliminaire (v0) de ThyroSeq utilisée comme <i>rule-in test</i> .
57	Li <i>et al.</i> , J Clin Endocrinol Metab 2011	L'étude économique porte sur l'évaluation d'une version préliminaire de l'analyse Afirma GEC [Chudova <i>et al.</i> , 2010] dont les caractéristiques de performance sont différentes de celles de la version finale [Alexander <i>et al.</i> , 2012]. De plus, le modèle utilisé a été développé à partir d'une version antérieure des recommandations de l'ATA (2009), qui diffère de celle présentement en vigueur (2015) quant à la prise en charge des nodules malins.
58	Nikiforov <i>et al.</i> , J Clin Endocrinol Metab 2009	Évaluation de la performance d'une version préliminaire (v0) de ThyroSeq utilisée comme <i>rule-in test</i> .

ElQ : écart interquartile; GEC : *gene expression classifier*; NIFTP : néoplasmes thyroïdiens folliculaires non invasifs avec caractéristiques nucléaires de type papillaire.

* Le nombre de nodules examinés est dit « limité » lorsque celui-ci est inférieur à 50 (seuil arbitraire établi en fonction du nombre de nodules évalués pour l'ensemble des études disponibles).

† Valeur attendue et non observée; l'estimation est basée sur le risque de malignité (prévalence de la maladie) des nodules à cytologie indéterminée évalué par les auteurs et les valeurs originales de sensibilité et de spécificité publiées antérieurement.

ANNEXE D

Résultats de l'évaluation de la qualité méthodologique des études

Tableau D-1 Évaluation de la qualité méthodologique des études de validation clinique

QUADAS-2 (version adaptée par Duh <i>et al.</i> , 2017)	ThyroSeqv2 Nikiforov 2014	ThyroSeqv2.1 Nikiforov 2015	ThyroSeqv3 Steward 2019	Afirma GEC Alexander 2012	Afirma GSC Patel 2018	ThyGenX/ ThyraMIR Labourier 2015
Sélection des patients						
La population à l'étude est-elle représentative des patients soumis à l'évaluation de première ligne d'un nodule thyroïdien à cytologie indéterminée?	Non-admissibilité basée sur la présence d'un nodule à cytologie de catégorie IV et sur la disponibilité d'un suivi chirurgical	Non-admissibilité basée sur la présence d'un nodule à cytologie de catégorie III et sur la disponibilité d'un suivi chirurgical	Non-admissibilité basée sur la disponibilité d'un suivi chirurgical (83 % nodules à CI)	Non-admissibilité basée sur la disponibilité d'un suivi chirurgical (72 % nodules à CI)	Non-admissibilité basée sur la disponibilité d'un suivi chirurgical (72 % nodules à CI)	Non-admissibilité basée sur la disponibilité d'un suivi chirurgical (40 % nodules à CI)
Les critères d'admissibilité pour l'évaluation moléculaire, y compris la proportion de nodules à cytologie indéterminée, sont-ils représentatifs de ceux observés dans la plupart des centres*?	Non/incertain	Non/incertain	Oui	Oui	Oui	Non/incertain
L'étude est-elle restreinte aux nodules appropriés (≥ 1 cm, catégories III/IV après une première biopsie)? Si d'autres catégories sont incluses, l'évaluation de la performance est-elle rapportée par sous-groupe?	Non/incertain	Non/incertain	Non/inclusion des nodules < 1 cm	Oui	Oui	Non/inclusion des nodules < 1 cm
Exécution du test						
Le test a-t-il été réalisé tel que recommandé?	Oui	Oui	Oui	Oui	Oui	Oui
Attribution d'un diagnostic standard de référence						
Un diagnostic standard de référence est-il attribué à la majorité des nodules inclus dans l'étude, moléculairement bénins et suspects?	Oui	Oui	Oui	Oui	Oui	Oui

QUADAS-2 (version adaptée par Duh <i>et al.</i>, 2017)	ThyroSeqv2 Nikiforov 2014	ThyroSeqv2.1 Nikiforov 2015	ThyroSeqv3 Steward 2019	Afirma GEC Alexander 2012	Afirma GSC Patel 2018	ThyGenX/ ThyraMIR Labourier 2015
Les individus qui ne sont pas soumis à une chirurgie immédiatement après le test moléculaire font-ils l'objet d'un suivi clinique?	s.o.	s.o.	s.o.	s.o.	s.o.	s.o.
Flux et chronologie						
Tous les patients reçoivent-ils un diagnostic standard de référence, histopathologie chirurgicale ou suivi clinique?	Oui	Oui	Oui	Oui	Oui	Oui
Le nombre de nodules exclus de l'analyse en raison de l'absence d'un diagnostic standard de référence est-il distribué uniformément parmi toutes les catégories de résultat (VP, FP, VN, FN)?	s.o.	s.o.	s.o.	s.o.	s.o.	s.o.
Aucune réattribution de catégories de résultat (VP, FP, VN, FN) n'a été effectuée sur la base de caractéristiques cliniques additionnelles	Oui	Oui	Oui	Oui	Oui	Oui

CI : cytologie indéterminée; FN : faux négatif; FP : faux positif; s.o. : sans objet; VN : vrai négatif; VP : vrai positif.

* Environ 20 % (min. 8 %; max. 36 %) des résultats de cytologie sont de catégorie indéterminée (III/IV) [Bongiovanni *et al.*, 2012].

Tableau D-2 Principales caractéristiques des études sélectionnées pour évaluer l'utilité clinique des tests moléculaires

Source	Période	Établissement	Devis	Conflit d'intérêts*	Nodules (cytologie/taille min. et max.)	Critères guidant l'usage du test
Afirma GEC						
Arosemena <i>et al.</i> , 2020	01/2012-01/2018	Jackson Memorial Hospital	Rétrospectif	Aucun	III : 100 % 0,9-3,1 cm	n.s.
San Martin <i>et al.</i> , 2020	12/2011-07/2017	Cleveland Clinic (Cleveland, Ohio)	Rétrospectif	Aucun	III/IV : 58/42 % 1,5-2,9 cm	n.s.
Sultan <i>et al.</i> , 2020	01/2014-10/2017	Henry Ford Hospital	Rétrospectif	n.s.	III/IV : 94/6 % Min. et max. n.s.	Emploi sélectif à la suite d'une 2 ^e BAAF de catégories III/IV
Endo <i>et al.</i> , 2019	02/2011-07/2017	Ohio State University Medical Center	Rétrospectif	n.s.	III/IV : 66/34 % 0,5-7,0 cm	n.s.
Angell <i>et al.</i> , 2019	01/2011-07/2017	Brigham and Women's Hospital Thyroid Nodule Clinic	Rétrospectif	Aucun	III/IV : 65/35 % 1,0-6,7 cm	Emploi sélectif basé sur la décision partagée en l'absence d'indication claire pour une chirurgie
Harrell <i>et al.</i> , 2019	01/2011-07/2017	Memorial Healthcare System	Rétrospectif	Aucun	III/IV : n.s. Min. et max. n.s.	n.s.
Wei <i>et al.</i> , 2019	07/2012-07/2017	Fox Chase Cancer Center	Rétrospectif	Oui (Veracyte)	III/IV : 80/20 % > 1 cm	n.s.
Deaver <i>et al.</i> , 2018	03/2011-12/2015	University of Colorado Hospital	Rétrospectif	Oui (Veracyte)	III/IV : 74/26 % Min. et max. n.s.	Emploi sélectif à la discrétion du clinicien
Livhits <i>et al.</i> , 2018	05/2016-06/2017	University of California Los Angeles Medical Center	Prospectif (randomisé)	Aucun	III/IV : 90/10 % 0 < taille > 4 cm	Emploi systématique
Kay-Rivest <i>et al.</i> , 2017	2013-2015	Jewish General Hospital; St. Clare's Mercy Hospital	Rétrospectif	Aucun	III/IV : 77/23 %; 33/67 % > 1 cm	Emploi sélectif basé sur la décision partagée
Yang <i>et al.</i> , 2016 [†]	08/2012-04/2014	University of California Los Angeles Medical Center	Rétrospectif	Aucun	III/IV : 87/13 % Min. et max. n.s.	Emploi systématique
Wu <i>et al.</i> , 2016	09/2012-02/2015	University of California Los Angeles Medical Center	Rétrospectif	Oui (Veracyte)	III/IV : 89/11 % 0 < taille > 4 cm	Emploi systématique
Villabona <i>et al.</i> , 2016	01/2012-01/2014	Cleveland Clinic Florida	Rétrospectif	Aucun	III : 100 % > 1 cm	Emploi sélectif à la discrétion du clinicien (1 ^{re} BAAF) ou à la suite d'une 2 ^e BAAF de catégorie III
Sipos <i>et al.</i> , 2016 [†]	09/2010-06/2014	Étude multicentrique (16 sites)	Rétrospectif	Oui (Veracyte)	III/IV : n.s. 1-7 cm	n.s.

Source	Période	Établissement	Devis	Conflit d'intérêts*	Nodules (cytologie/taille min. et max.)	Critères guidant l'usage du test
Singer <i>et al.</i> , 2016 [‡]	01/2011-08/2013	Étude multicentrique	Rétrospectif	Oui (Veracyte)	III/IV : n.s. ≥ 1 cm	n.s.
Samulski <i>et al.</i> , 2016	01/2011-12/2015	Hospital of the University of Pennsylvania	Rétrospectif	Oui (Veracyte)	III/IV : 57/43 % 0 < taille > 4 cm	Emploi sélectif à la discrétion du clinicien
Sacks <i>et al.</i> , 2016	07/2013-12/2014	Cedars-Sinai Medical Center	Rétrospectif	Aucun	III/IV : n.s. valeur ambiguë	Emploi sélectif à la discrétion du clinicien
Chaudhary <i>et al.</i> , 2016	07/2012-06/2015	Ohio State University Wexner Medical Center	Rétrospectif	Aucun	III/IV : 46/44 % Min. et max. n.s.	Emploi sélectif à la discrétion du clinicien
Marti <i>et al.</i> , 2015	02/2013-12/2014	Memorial Sloan Kettering Cancer Center; Mount Sinai Beth Israel	Rétrospectif	Aucun	III/IV : 67/33 %; 56/44 % Min. et max. n.s.	n.s.
Angell <i>et al.</i> , 2015 [‡]	11/2010-04/2014	Brigham and Women's Hospital Thyroid Nodule Clinic	Rétrospectif	Aucun	III/IV : 54/46 % 1,0-5,3 cm	Emploi sélectif à la discrétion du clinicien
Mclver <i>et al.</i> , 2014	05/2011-12/2012	Mayo Clinic (Rochester/Jacksonville)	Prospectif	Oui (Veracyte et Asuragen)	III/IV : 8/92 % > 1 cm	Emploi sélectif basé sur la décision partagée en l'absence d'indication claire pour une chirurgie
Afirma GSC						
San Martin <i>et al.</i> , 2020	07/2017-09/2018	Cleveland Clinic (Cleveland, Ohio)	Rétrospectif	Aucun	III/IV : 63/37 % 1,4-3,2 cm	n.s.
Endo <i>et al.</i> , 2019	07/2017-12/2018	Ohio State University Medical Center	Rétrospectif	n.s.	III/IV : 76/24 % 0,65-8,9 cm	n.s.
Angell <i>et al.</i> , 2019	07/2017-08/2018	Brigham and Women's Hospital Thyroid Nodule Clinic	Rétrospectif	Aucun	III/IV : 59/41 % 1,0-6,9 cm	Emploi sélectif basé sur la décision partagée en l'absence d'indication claire pour une chirurgie
Harrell <i>et al.</i> , 2019	08/2017-07/2018	Memorial Healthcare System	Rétrospectif	Aucun	III/IV : n.s. Min. et max. n.s.	n.s.
Wei <i>et al.</i> , 2019	07/2017-01/2019	Fox Chase Cancer Center	Rétrospectif	Oui (Veracyte)	III/IV : 85/15 % > 1 cm	n.s.
ThyroSeq v2.1						
Livhits <i>et al.</i> , 2018	05/2016-06/2017	University of California Los Angeles Medical Center	Prospectif	Aucun	III/IV : 90/10 % 0 < taille > 4 cm	Emploi systématique
Taye <i>et al.</i> , 2018	11/2014-04/2016	Mount Sinai Beth Israel; Mount Sinai West & St. Luke's Hospital; Mount Sinai Hospital	Rétrospectif	n.s.	III/IV : 79/21 % 0,5-7,2 cm	Emploi sélectif ou systématique selon le médecin prescripteur
Valderrabano <i>et al.</i> , 2017	09/2014-04/2016	Moffitt Cancer Center and Research Institute	Rétrospectif	Oui (Rosetta Genomics)	III/IV : 55/45 % 0 ^s < taille > 4 cm	Emploi systématique

Source	Période	Établissement	Devis	Conflit d'intérêts*	Nodules (cytologie/taille min. et max.)	Critères guidant l'usage du test
ThyroSeq v3						
Chen <i>et al.</i> , 2020	01/2019-02/2019	Royal Victoria Hospital	Rétrospectif	Aucun	III/IV : 42/58 % 1,1-4,2 cm	Emploi sélectif

n.s. : non spécifié.

* Conflits en lien avec les tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée.

† La cohorte de patients du groupe GEC de cette étude chevauche celle décrite par Wu et ses collaborateurs [2016]. Néanmoins, l'étude menée par Yang et ses collaborateurs [2016] présente une comparaison de la prise en charge des nodules à cytologie indéterminée selon la présence ou l'absence d'une évaluation moléculaire.

‡ Les données présentées sont spécifiques aux nodules GEC bénins.

§ Trois nodules mesuraient ≤ 1 cm; aucun n'avait eu de suivi chirurgical.

Tableau D-3 Évaluation de la qualité des études économiques (selon l'outil CASP)

	Zanocco <i>et al.</i> , 2020	Nicholson <i>et al.</i> , 2019	Balentine <i>et al.</i> , 2018
A. Cette évaluation économique est-elle valide?			
1- L'évaluation repose-t-elle sur une question bien définie?	Oui	Oui	Oui
2- A-t-on fait une description complète des options comparées?	Oui	Oui	Oui
3- L'article démontre-t-il l'efficacité de l'intervention? (l'innocuité de l'intervention est-elle acceptable?)	Oui	Oui	Oui
4- Les effets de l'intervention ont-ils été identifiés, mesurés et évalués adéquatement?	QALY	Bon Dx	QALY
B. Comment a-t-on évalué et comparé les coûts et les conséquences?			
5- A-t-on identifié, mesuré avec les unités appropriées et évalué de façon vraisemblable toutes les ressources et tous les coûts importants et pertinents pour chaque option considérée?	Non	Non	Non
6- Les auteurs ont-ils ajusté les coûts et les conséquences en fonction du moment où ils se concrétiseront (actualisation)?	Oui	Partiellement (les conséquences n'ont pas été actualisées)	Oui
7- Quels sont les résultats de l'évaluation?	Afirma GEC : efficient en l'absence de facteurs de risque échographique élevé	Coût/bon Dx inférieur avec ThyroSeq v3 ou Afirma GSC vs lobectomie	Afirma GEC : non efficient
8- Une analyse différentielle des conséquences et du coût a-t-elle été réalisée pour les options comparées?	Oui	Non	Oui
9- Une analyse de sensibilité en bonne et due forme a-t-elle été effectuée?	Oui	Oui	Oui
C. Les résultats vont-ils faciliter l'achat de services pour la population locale?			
10- Le modèle aura-t-il la même efficacité dans le contexte québécois?	Non	Difficile à dire	Non
11- Les coûts sont-ils transposables au contexte québécois?	Non	Non	Non
12- L'utilisation du même modèle dans le contexte québécois est-elle justifiée?	Non	Non	Non

Dx : diagnostic; QALY : *quality-adjusted life year*.

ANNEXE E

Impact de la prévalence de la maladie

Tableau E-1 Variation de la valeur prédictive négative et du risque résiduel selon la prévalence locale de la maladie sur différents sites des études de validation clinique

Afirma GSC (sensibilité : 91 %; spécificité : 68 %)				
Validation clinique (43 sites)	Prévalence maladie : 24 % Probabilité d'un faux négatif : $9 \% \times 24 \% = 0,0216$	Prévalence nodules bénins : 76 % Probabilité d'un vrai négatif : $68 \% \times 76 \% = 0,5168$	VPN [0,5168 / (0,5168 + 0,0216)] 96 %	Risque résiduel 4 %
SITE n° 3	Prévalence maladie : 56 % Probabilité d'un faux négatif : $9 \% \times 56 \% = 0,0504$	Prévalence nodules bénins : 44 % Probabilité d'un vrai négatif : $68 \% \times 44 \% = 0,2992$	VPN [0,2992 / (0,2992 + 0,0504)] 86 %	Risque résiduel 14 %
SITE n° 32	Prévalence maladie : 19 % Probabilité d'un faux négatif : $9 \% \times 19 \% = 0,0171$	Prévalence nodules bénins : 81 % Probabilité d'un vrai négatif : $68 \% \times 81 \% = 0,5508$	VPN [0,5508 / (0,5508 + 0,0171)] 97 %	Risque résiduel 3 %
ThyroSeq v3 (sensibilité : 94 %; spécificité : 82 %)				
Validation clinique (10 sites)	Prévalence maladie : 28 % Probabilité d'un faux négatif : $6 \% \times 28 \% = 0,0168$	Prévalence nodules bénins : 72 % Probabilité d'un vrai négatif : $82 \% \times 72 \% = 0,5904$	VPN [0,5904 / (0,5904 + 0,0168)] 97 %	Risque résiduel 3 %
SITE n° 4	Prévalence maladie : 60 % Probabilité d'un faux négatif : $6 \% \times 60 \% = 0,0360$	Prévalence nodules bénins : 40 % Probabilité d'un vrai négatif : $82 \% \times 40 \% = 0,3280$	VPN [0,3280 / (0,3280 + 0,0360)] 90 %	Risque résiduel 10 %
SITE n° 8	Prévalence maladie : 9 % Probabilité d'un faux négatif : $6 \% \times 9 \% = 0,0054$	Prévalence nodules bénins : 91 % Probabilité d'un vrai négatif : $82 \% \times 91 \% = 0,7462$	VPN [0,7462 / (0,7462 + 0,0054)] 99 %	Risque résiduel 1 %

VPN : valeur prédictive négative; [vrais négatifs ÷ (vrais négatifs + faux négatifs)].

Validité analytique

Tableau E-2 Caractéristiques de performance analytique des tests Afirma GSC et ThyroSeq v3

Afirma GSC [Hao <i>et al.</i> , 2019]	Sensibilité et spécificité analytique		
	Contenu minimal en ARN* : 5 ng		
	Contenu minimal en cellules malignes : 25 %		
	Tolérance à la contamination en contenu sanguin : 75 %		
	Tolérance à la contamination en ADN génomique : 30 %		
	Répétabilité et reproductibilité		
		Écart-type des scores GSC (IC 95 %)[†]	
	Répétabilité (<i>intra-run</i>)	0,069 (0,053-0,073)	
	Reproductibilité (<i>inter-run</i>)	0,274 (0,228-0,310)	

ThyroSeq v3 [Nikiforova <i>et al.</i> , 2018]	Sensibilité et spécificité analytique		
	Contenu minimal en acides nucléiques (ADN et ARN) [‡] : 2,5 ng		
	Contenu minimal en cellules malignes : 12 %		
	Tolérance à la contamination en contenu sanguin : 88 %		
	Répétabilité et reproductibilité		
		Fréquence allélique (ADN)	Niveau d'expression (ARN)
	Répétabilité (<i>intra-run</i>)	0,30-1,1 %	0,29-0,69 %
	Reproductibilité (<i>inter-run</i>)	0,26-1,42 %	0,21-1,49 %

ADN : acide désoxyribonucléique; ARN : acide ribonucléique; GSC : *genomic sequencing classifier*.

* Les auteurs mentionnent que le fabricant recommande l'utilisation de 10 ng d'acides nucléiques (ADN et ARN); pour l'évaluation de la sensibilité analytique, entre 5 et 30 ng d'acides nucléiques ont été utilisés.

[†] Le seuil de variation toléré (sans impact sur la sensibilité et la spécificité) correspond à un écart-type de 0,44.

[‡] Le contenu standard en ARN est de 15 ng; pour l'évaluation de la sensibilité analytique, entre 0,5 et 10 ng d'ARN ont été utilisés.

ANNEXE F

Estimation de la population admissible, analyses de scénarios pour l'impact budgétaire et estimation des coûts unitaires

Tableau F-1 Estimation de la population admissible (ThyroSeq v3 vs lobectomie)

	AN 1	AN 2	AN3	TOTAL
Statu quo : lobectomie				
Nombre de lobectomies	969	969	969	2 907
Nombre d'hypothyroïdies post-lobectomie	213	426	639	639
Nouveau scénario : option ThyroSeq v3				
Nombre de tests	969	969	969	2 907
Nombre de nodules malins	242	242	242	726
Nombre de nodules bénins	727	727	727	2 181
Nombre de nodules malins vrais positifs	228	228	228	684
Nombre de nodules malins faux négatifs	15	15	15	45
Nombre de nodules bénins vrais négatifs	596	596	596	1 788
Nombre de nodules bénins faux positifs	131	131	131	393
Nombre de nodules malins vrais positifs réséqués	175	175	175	525
Nombre de nodules malins faux négatifs réséqués	2	2	2	6
Nombre total de nodules malins réséqués	177	177	177	531
Nombre de nodules malins non réséqués	65	130	195	195
Nombre de nodules bénins faux positifs réséqués	101	101	101	303
Nombre de nodules bénins négatifs réséqués	77	77	77	231
Nombre total de nodules bénins réséqués	178	178	178	534
Nombre total de lobectomies	356	356	356	1 068
Nombre de nodules sous surveillance	613	1 226	1 839	1 839
Nombre d'hypothyroïdies post-lobectomie	78	156	234	234

Tableau F-2 Estimation de la population admissible (Afirma GSC vs lobectomie)

	AN 1	AN 2	AN3	TOTAL
Statu quo : lobectomie				
Nombre de lobectomies	969	969	969	2 907
Nombre d'hypothyroïdies post-lobectomie	213	426	639	639
Nouveau scénario : option Afirma GSC				
Nombre de tests	969	969	969	2 907
Nombre de nodules malins	242	242	242	727
Nombre de nodules bénins	727	727	727	2 181
Nombre de nodules malins vrais positifs	221	221	221	662
Nombre de nodules malins faux négatifs	22	22	22	66
Nombre de nodules bénins vrais négatifs	494	494	494	1 482
Nombre de nodules bénins faux positifs	233	233	233	699
Nombre de nodules malins vrais positifs réséqués	170	170	170	510
Nombre de nodules malins faux négatifs réséqués	3	3	3	9
Nombre total de nodules malins réséqués	173	173	173	519
Nombre de nodules malins non réséqués	70	140	210	210
Nombre de nodules bénins faux positifs réséqués	179	179	179	537
Nombre de nodules bénins négatifs réséqués	64	64	64	192
Nombre total de nodules bénins réséqués	243	243	243	729
Nombre total de lobectomies	416	416	416	1 248
Nombre de nodules sous surveillance	553	1 106	1 659	1 659
Nombre d'hypothyroïdies post-lobectomie	92	184	276	276

Tableau F-3 Analyses de scénarios pour l'impact budgétaire (ThyroSeq v3 vs lobectomie)

Scénario	Paramètre évalué	Impact net	Modification par rapport au scénario de base
Scénario de base		- 960 473 \$	0 %
Nombre de chirurgies pour nodules thyroïdiens au Québec	3 000	- 1 166 706 \$	21 %
	2 000	- 765 474 \$	- 20 %
Proportion de nodules à cytologie indéterminée réséqués	0,6	- 1 348 374 \$	40 %
	0,2	- 424 799 \$	- 56 %
Risque de malignité pour un nodule à cytologie indéterminée	0,4	947 001 \$	- 199 %
	0,1	- 2 867 946 \$	199 %
Sensibilité de ThyroSeq v3	0,98	- 793 151 \$	- 17 %
	0,86	- 1 295 117 \$	35 %
Spécificité de ThyroSeq v3	0,87	- 1 587 931 \$	65 %
	0,75	- 82 031 \$	- 91 %
Risque de lobectomie si le résultat ThyroSeq v3 est suspect	0,97	974 191 \$	- 201 %
	0,54	- 3 185 336 \$	232 %
Risque de lobectomie si le résultat ThyroSeq v3 est bénin	0,48	4 804 301 \$	- 600 %
	0	- 3 101 674 \$	223 %
Risque d'hypothyroïdie post-lobectomie	0,27	- 998 006 \$	3,9 %
	0,19	- 937 953 \$	- 2,3 %
Coût du test	6 731 \$	2 953 325 \$	- 407,5 %
	4 039 \$	- 4 874 271 \$	407,5 %

Tableau F-4 Analyses de scénarios pour l'impact budgétaire (Afirma GSC vs lobectomie)

Scénario	Paramètre évalué	Impact net	Modification par rapport au scénario de base
Scénario de base		677 330 \$	0 %
Nombre de chirurgies pour nodules thyroïdiens au Québec	3 000	817 374 \$	21 %
	2 000	544 916 \$	- 20 %
Proportion de nodules à cytologie indéterminée réséqués	0,6	940 736 \$	39 %
	0,2	313 579 \$	- 54 %
Risque de malignité pour un nodule à cytologie indéterminée	0,4	2 163 859 \$	219 %
	0,1	- 809 199 \$	- 219 %
Sensibilité d'Afirma GSC	0,98	971 277 \$	43 %
	0,79	173 422 \$	- 74 %
Spécificité d'Afirma GSC	0,76	- 330 486 \$	- 149 %
	0,60	1 685 146 \$	149 %
Risque de lobectomie si le résultat Afirma GSC est suspect	0,97	3 131 258 \$	362 %
	0,54	- 2 144 687 \$	- 417 %
Risque de lobectomie si le résultat Afirma GSC est bénin	0,48	5 568 782 \$	722 %
	0	- 1 139 495 \$	- 268 %
Risque d'hypothyroïdie post-lobectomie	0,27	651 095 \$	- 3,9 %
	0,19	693 071 \$	2,3 %
Coût du test	6 731 \$	4 591 128 \$	577,8 %
	4 039 \$	- 3 236 468 \$	- 577,8 %

Tableau F-5 Estimation des coûts unitaires

Actes	Code	Item et unités	Coût unitaire	Coût total	Référence
Opération chirurgicale (lobectomie)	06152	Honoraires du chirurgien	584,25 \$	584,25 \$	RAMQ <i>Manuel des spécialistes</i>
		Honoraires de l'anesthésiste (6 unités)	16,96 \$	101,76 \$	RAMQ <i>Manuel des spécialistes</i>
Frais d'établissement	CA 6260	Bloc opératoire (chirurgie majeure III - Lourdeur 1,270)	3 823,00 \$	3 823,00 \$	MSSS Programme d'accès à la chirurgie
	APR-DRG (404)	Hospitalisation (NIRRU-cas = 1,18)	4 039,57 \$	4 766,69 \$	MSSS APR-DRG
Visite de contrôle en hospitalisation	09152	Honoraires du médecin (2 visites)	19,30 \$	38,60 \$	RAMQ <i>Manuel des spécialistes</i>
Coût total lobectomie				9 314,30 \$	
Échographie de suivi du nodule	08330	Honoraires du médecin	28,75 \$	28,75 \$	RAMQ <i>Manuel des spécialistes</i>
Visite médicale de suivi nodule	15391	Honoraires du médecin	42,20 \$	42,20 \$	RAMQ <i>Manuel des spécialistes</i>
Frais d'établissements – service de consultations externes – spécialisés externes			53,01 \$	53,01 \$	MSSS Coût du NIRRU
Frais d'établissements échographie	8330	15 unités techniques provinciales	3,65 \$	54,80 \$	MSSS Valeurs unitaires
Coût total surveillance du nodule				178,76 \$	
Coût annuel de l'hormonothérapie de substitution (hypothyroïdie) : Synthroid 0,025 mg + 6,5 % de marge du bénéficiaire et honoraires de pharmacien				62,86 \$	RAMQ <i>Manuel des pharmaciens et Liste des médicaments</i>
Coût total du traitement et de la surveillance de l'hypothyroïdie (visite médicale, frais d'établissement + hormonothérapie de substitution)				158,07\$	

APR-DRG : All Patient Refined - Diagnosis Related Groups; NIRRU : niveau d'intensité relative des ressources utilisées.

ANNEXE G

Gradation de la preuve

Tableau G-1 Appréciation de la preuve scientifique

PARAMÈTRES	CRITÈRES D'APPRÉCIATION DE LA PREUVE			NIVEAU DE PREUVE
	QUALITÉ	COHÉRENCE	TRANSFÉRABILITÉ	
Validité clinique - versions disponibles	ThyroSeq v3 (1 étude) Afirma GSC (1 étude) ThyGeNEXT/ThyraMIR (aucune étude) De faible à modérée Une seule étude de validation clinique par test dont les plans sont appropriés pour évaluer la performance; le risque de biais est modéré.	Sans objet La performance des tests dans les études de validation clinique originales n'a pas été confirmée par des études indépendantes.	Incertaine La population étudiée pourrait différer de la population ciblée. Par conséquent, les résultats rapportés dans la littérature ne sont pas directement généralisables à la population ciblée.	Faible La plupart des critères ont obtenu une appréciation négative. Cependant, il est probable que l'application de certaines mesures permette de générer, dans la population ciblée, des caractéristiques de performance similaires à celles observées dans les études de validation clinique originales.
Utilité clinique - taux de chirurgies	Afirma GEC (20 études) Afirma GSC (5 études) ThyroSeqv2.1 (3 études) ThyroSeq v3 (1 étude) Faible Plusieurs études dont les plans sont peu appropriés pour répondre à la question d'évaluation; le risque de biais est élevé.	Afirma GEC : modérée/élevée Afirma GSC : élevée ThyroSeqv2.1 : modérée ThyroSeq v3 : sans objet La plupart des études sont cohérentes et l'incohérence peut généralement être expliquée	Afirma GEC : aucune ThyroSeqv2.1 : aucune Ces versions ne sont plus disponibles. Afirma GSC : incertaine ThyroSeq v3 : incertaine La population étudiée (caractéristiques des nodules et critères guidant l'usage du test) pourrait différer de la population ciblée. Par conséquent, les résultats rapportés dans la littérature ne sont pas directement généralisables à la population ciblée.	Faible La plupart des critères ont obtenu une appréciation négative. Cependant, il est probable qu'un usage sélectif des tests moléculaires basé sur les caractéristiques appropriées permette de réduire le taux de chirurgies.

PARAMÈTRES	CRITÈRES D'APPRÉCIATION DE LA PREUVE			NIVEAU DE PREUVE
	QUALITÉ	COHÉRENCE	TRANSFÉRABILITÉ	
- risque résiduel de malignité des nodules mol. bénins	Afirma GEC (1 étude) Faible Une seule étude portant sur peu de nodules GEC bénins dont les plans sont plus ou moins appropriés pour répondre à la question d'évaluation; le risque de biais est modéré/élevé.	Sans objet (une seule étude)	Aucune Afirma GEC n'est plus disponible.	Insatisfaisant Les données disponibles sont insuffisantes.

GEC : *gene expression classifier*; GSC : *genomic sequencing classifier*.



Siège social

2535, boulevard Laurier, 5^e étage
Québec (Québec) G1V 4M3
418 643-1339

Bureau de Montréal

2021, avenue Union, 12^e étage, bureau 1200
Montréal (Québec) H3A 2S9
514 873-2563

inesss.qc.ca

**Institut national
d'excellence en santé
et en services sociaux**

Québec 

