

Validation analytique de tests cliniques
effectués par séquençage de nouvelle
génération (SNG) en génétique
humaine

Recommandations de bonnes pratiques

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

Direction de l'évaluation des médicaments
et des technologies à des fins de
remboursement



Validation analytique de tests cliniques effectués par séquençage de nouvelle génération (SNG) en génétique humaine

Recommandations de bonnes pratiques

Rédaction

Catherine Gravel

Collaboration

Frédéric Breton

Coordination scientifique

Éric Potvin

Direction

Michèle de Guise



Le contenu de ce rapport a été rédigé et édité par l'Institut national d'excellence en santé et en services sociaux (INESSS).

Membres de l'équipe de projet

Auteure principale

Catherine Gravel, M. Sc., D.E.S.S.

Coordonnateur scientifique

Éric Potvin, Ph. D.

Collaborateur interne

Frédéric Breton, B. Sc.

Directrice

Michèle de Guise, M.D., M.M, FRCPC

Repérage d'information scientifique

Renaud Lussier, M.S.I.

Lysane St-Amour, M.B.S.I.

Soutien administratif

Christine Lemire

Équipe de l'édition

Denis Santerre
Hélène St-Hilaire
Nathalie Vanier

Sous la coordination de

Renée Latulippe, M.A.

Avec la collaboration de

Gilles Bordage, révision linguistique

Dépôt légal

Bibliothèque et Archives nationales du Québec, 2021

Bibliothèque et Archives Canada, 2021

ISBN 978-2-550-88550-4 (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). Validation analytique de tests cliniques effectués par séquençage de nouvelle génération (SNG) en génétique humaine – Recommandations de bonnes pratiques. Rapport rédigé par Catherine Gravel. Québec, Qc : INESSS; 2021. 70 p.

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

Comité consultatif

Les membres du comité d'experts pour ce rapport sont :

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

D^r Guy Fink, biochimiste clinique, CIUSSS de l'Estrie – CHU de Sherbrooke

D^r Louis Gaboury, anatomopathologiste, Centre hospitalier de l'Université de Montréal

D^r David Rosenblatt, pédiatre, généticien, Centre universitaire de santé McGill

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

Lecteurs externes

La lecture externe est un des mécanismes utilisés par l'INESSS pour assurer la qualité de ses travaux. Les lecteurs externes valident, entre autres choses, l'exactitude du contenu en fonction de leur domaine d'expertise respectif. Pour ce rapport, les lecteurs externes sont :

D^r Jonatan Blais, médecin biochimiste, Centre intégré de santé et de services sociaux (CISSS) de Chaudière-Appalaches

D^{re} Isabelle De Bie, généticienne, Centre universitaire de santé McGill

D^r Patrice Desmeules, anatomopathologiste, Institut universitaire de cardiologie et de pneumologie de Québec

Comité scientifique permanent des analyses de biologie médicale

Les membres du Comité scientifique permanent des analyses de biologie médicale (CSABM) de l'INESSS sont :

Président

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

Vice-président

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

Membres

M^{me} Suzanne K. Bédard, membre citoyenne

M^{me} Lorraine Caron, consultante en éthique

D^r Guy Fink, biochimiste clinique, CIUSSS de l'Estrie – CHU de Sherbrooke

D^r David Rosenblatt, pédiatre, généticien, Centre universitaire de santé McGill

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

Déclaration de conflit d'intérêts

L'INESSS ne déclare aucun conflit d'intérêts.

Responsabilité

L'Institut assume l'entière responsabilité de la forme et du contenu définitifs de ce document. Les conclusions et recommandations de ce rapport ne reflètent pas forcément les opinions des lecteurs externes ou des autres personnes consultées dans le cadre de ces travaux.

TABLE DES MATIÈRES

SIGLES ET ABRÉVIATIONS.....	I
NOTE AUX LECTEURS	II
1. INTRODUCTION.....	1
1.1. Contexte de l’amorce des travaux.....	1
1.2. Objectifs	1
1.3. Aspects exclus.....	2
2. MÉTHODOLOGIE.....	3
2.1. Question d’évaluation	3
2.2. Revue de littérature	3
2.3. Processus de participation	3
3. PRÉAMBULE.....	4
4. OBJECTIF DE PERFORMANCE RECOMMANDÉ PAR LE GROUPE DE TRAVAIL.....	7
5. UTILISATION CLINIQUE ET TYPE D’ANALYSE.....	8
5.1. Utilisation clinique du SNG.....	8
5.2. Type d’analyse	9
6. PLANIFICATION PRÉANALYTIQUE.....	10
6.1. Détermination des séquences de référence	11
6.2. Conception, validation et choix des amorces.....	12
6.3. Détermination des types de spécimens cliniques acceptés pour un test effectué par SNG.....	13
6.4. Détermination du matériel de référence et des échantillons cliniques témoins nécessaires à la validation analytique d’un test effectué par SNG.....	13
7. VALIDATION ANALYTIQUE.....	16
7.1. Particularités du SNG	16
7.1.1. Validation de la méthode de SNG	17
7.1.2. Validation du pipeline informatique	18
7.1.3. Validation d’un test	19
7.2. Évaluation de la qualité des données obtenues par SNG	19
7.2.1. Valeur de qualité de l’appel des bases	22
7.2.2. Couverture de la séquence ciblée	22
7.3. Taux d’erreur des technologies de SNG.....	25
7.4. Critères de performance à évaluer lors de la validation analytique d’un test effectué par SNG	27
7.5. Détermination du nombre de variations nucléotidiques et d’échantillons cliniques à analyser pour l’évaluation de la validité analytique d’un test effectué par SNG	29
7.6. Estimation de la puissance statistique d’une étude de validation analytique en fonction de la taille de l’échantillon	33
8. ÉTAPES POSTANALYTIQUES.....	36
8.1. Interprétation des variations nucléotidiques héréditaires et communication des données de séquence	36

8.2. Interprétation des variations nucléotidiques somatiques	38
8.3. Assurance de la qualité et épreuves de compétence	41
8.3.1. Contrôles internes de la qualité	41
8.3.2. Contrôles externes de la qualité	41
RÉFÉRENCES	43
ANNEXE A	48
Stratégies de repérage d'information scientifique	48
ANNEXE B	50
Description du principe technologique du SNG et précisions techniques	50
ANNEXE C	52
Paramètres et mesures de qualité du SNG	52
ANNEXE D	54
Description des critères de performance nécessaires à la validation de la méthode ou d'un test effectué par séquençage de nouvelle génération (SNG)	54
ANNEXE E	59
Recommandations d'organismes professionnels et réglementaires concernant les critères de performance à évaluer lors de la validation analytique d'un test clinique effectué par SNG ou dans le contexte des analyses génétiques et moléculaires	59
ANNEXE F	60
Validation analytique d'un test clinique effectué par séquençage de nouvelle génération (SNG) : recommandations de bonnes pratiques	60

LISTE DES TABLEAUX

Tableau 1	Tests conçus en laboratoire (TCL).....	4
Tableau 2	Résultats et profils génétiques pouvant être obtenus à la suite de l'analyse de l'ADN par séquençage	9
Tableau 3	Recommandations du CLSI concernant le matériel de référence et les échantillons cliniques à utiliser lors de la validation analytique d'un test de génétique moléculaire	14
Tableau 4	Critères nécessaires à l'évaluation des performances analytiques d'un essai de SNG	20
Tableau 5	Correspondance entre la valeur de qualité et la probabilité d'erreur de l'appel de la base à une position donnée de la séquence	22
Tableau 6	Recommandations d'organismes professionnels et réglementaires concernant les valeurs de qualité minimales de l'appel des bases et de profondeur de lecture pour l'appel d'une variation nucléotidique	24
Tableau 7	Fréquence et types d'erreurs de différentes technologies et plateformes de séquençage....	26
Tableau 8	Définitions des critères de performance à évaluer lors de la procédure de validation analytique d'un test clinique effectué par SNG.....	27
Tableau 9	Recommandations d'organismes professionnels et réglementaires concernant le nombre minimal de variations nucléotidiques ou d'échantillons cliniques requis pour la validation analytique d'un test clinique par SNG	31
Tableau 10	Effet de la taille de l'échantillon sur la puissance statistique lors de la détermination de la sensibilité d'un test qualitatif	33
Tableau 11	Catégorisation des variations nucléotidiques somatiques en fonction du niveau de preuve scientifique	39

LISTE DES FIGURES

Figure 1	Processus de mise en place d'un nouveau test de génétique moléculaire aux fins d'une utilisation en diagnostic médical	5
Figure 2	Incertitude de l'estimation d'une proportion de 100 % en fonction du nombre d'observations	34

SIGLES ET ABRÉVIATIONS

ACGS	Association for Clinical Genetic Science
ACMG	American College of Medical Genetics and Genomics
ADN	Acide désoxyribonucléique
AMP	Association for Molecular Pathology
ANPGM	Association nationale des praticiens de génétique moléculaire
CAP	College of American Pathologists
CCGM	Collège canadien des généticiens médicaux (ou CCMG, de l'anglais Canadian College of Medical Geneticists)
CG	Cytosines et guanines
CLSI	Clinical and Laboratory Standards Institute
CSA	Canadian Standards Association
CSABM	Comité scientifique permanent des analyses de biologie médicale
DDIV	Dispositif de diagnostic <i>in vitro</i>
FDA	Food and Drug Administration
FFEP	Fixés dans la formaline et enrobés de paraffine (ou FFPE, de l'anglais <i>formalin-fixed, paraffin-embedded</i>)
INCa	Institut national du cancer
INESSS	Institut national d'excellence en santé et en services sociaux
IQN Path	International Quality Network for Pathology
ISO	Organisation internationale de normalisation
LRG	<i>Locus Reference Genomic</i>
MSSS	Ministère de la Santé et des Services sociaux
NCBI	National Center for Biotechnology Information
Nex-StoCT	Groupe de travail <i>Next-generation Sequencing: Standardization of Clinical Testing</i>
NIH	National Institute of Health
SNG	Séquençage de nouvelle génération
TAAN	Technique d'amplification des acides nucléiques
TCL	Test conçu en laboratoire (ou LDT, de l'anglais <i>Laboratory Developed Test</i>)
VMN	Variations de plusieurs nucléotides (ou MNV, de l'anglais <i>multiple nucleotides variants</i>)
VNC	Variations du nombre de copies (ou CNV, de l'anglais <i>copy number variants</i>)
VSN	Variations d'un seul nucléotide (ou SNV, de l'anglais <i>single nucleotide variant</i>)

NOTE AUX LECTEURS

Le présent document est fondé sur les recommandations d'organismes professionnels et réglementaires ainsi que sur une recherche documentaire complémentaire selon les données disponibles au moment de sa production par l'INESSS.

Ce document ne doit pas être considéré comme un recueil de normes ou d'exigences servant à accepter ou refuser un dossier de validation analytique aux fins de désignation d'un laboratoire pour une analyse donnée. Il s'agit de recommandations de bonnes pratiques en médecine de laboratoire appliquées aux analyses cliniques effectuées par séquençage de nouvelle génération.

L'INESSS est conscient que les recommandations incluses dans ce rapport ne sont pas applicables à toutes les situations. Il peut exister différents types de contraintes pouvant empêcher un laboratoire, pour une analyse donnée, de réaliser toutes les mesures proposées. Le nombre d'échantillons cliniques à inclure dans une étude de validation analytique de puissance statistique optimale est une de ces contraintes.

1. INTRODUCTION

1.1. Contexte de l'amorce des travaux

À la demande du ministère de la Santé et des Services sociaux (MSSS) du Québec, l'Institut national d'excellence en santé et en services sociaux (INESSS) a élaboré, en 2012, un processus d'évaluation des analyses de biologie médicale dont le résultat final est la transmission d'un avis au ministre de la Santé et des Services sociaux. Cette évaluation couvre différents volets, dont celui de la validité analytique. L'évaluation du volet analytique repose essentiellement sur une appréciation du plan de validation proposé par les laboratoires demandeurs et, le cas échéant, propose certaines améliorations, l'objectif étant d'aider chaque laboratoire à produire les données de validation analytique nécessaires pour garantir au médecin prescripteur et au patient la fiabilité des résultats. En juin 2018, le MSSS a demandé aux établissements du réseau d'obtenir l'accréditation ISO 15189, qui a son propre cahier des charges en matière de validation et de vérification. Les normes ISO exigent qu'une validation soit aussi étendue que nécessaire et confirme, par des preuves tangibles, que les exigences spécifiques pour l'utilisation prévue de l'examen ont été satisfaites.

Dans ce contexte, il a été convenu de produire un document destiné à guider les laboratoires qui devront effectuer la validation analytique d'une analyse de biologie médicale réalisée par séquençage de nouvelle génération (SNG).

1.2. Objectifs

Ce document traite de la validation analytique de tests cliniques effectués par séquençage de nouvelle génération (SNG) appliqués aux conditions et maladies héréditaires ou acquises dans le cadre d'un processus néoplasique. Il concerne les applications spécifiques à la caractérisation de régions définies du génome, soit l'analyse d'un ou plusieurs gènes ou de régions de gènes, pour lesquels il est nécessaire de procéder à une validation analytique complète du test.

Les principaux objectifs de ce rapport sont les suivants :

- 1) Résumer les paramètres et éléments à évaluer et à documenter lors de la validation analytique d'un nouveau test effectué par SNG et conçu en laboratoire ou effectué à l'aide d'une trousse non homologuée pour un usage clinique (voir le [tableau 1](#) concernant la définition des tests conçus en laboratoire);
- 2) Proposer, lorsque cela est pertinent, un devis expérimental permettant de mesurer les paramètres de performance analytique d'un test, et ce, avec une exactitude statistiquement satisfaisante pour garantir un taux d'erreur inférieur ou égal à un seuil prédéfini. De ce fait, certaines recommandations concernant le nombre d'échantillons requis pour effectuer la validation analytique de la méthode de séquençage ou d'un test réalisé par SNG sont proposées.

Un résumé de certaines normes et recommandations relatives aux différentes étapes des processus préanalytiques et postanalytiques est également présenté.

Comme il a été mentionné précédemment, le présent document ne doit pas être considéré comme un recueil de normes ou d'exigences servant à accepter ou refuser un dossier de validation analytique aux fins de désignation d'un laboratoire pour une analyse donnée. Les membres du CSABM sont conscients qu'il existe différents types de contraintes pouvant empêcher un laboratoire, pour une analyse donnée, de réaliser toutes les mesures proposées dans ce document et qui seront justifiées dans le rapport de validation analytique.

1.3. Aspects exclus

Le présent document ne concerne pas :

- 1) la validation analytique de tests effectués par SNG dans le domaine de la microbiologie ou de l'infectiologie;
- 2) la vérification analytique de tests dont les paramètres de performance analytique sont établis et disponibles (ex. : trousse commerciale homologuée par Santé Canada ou la Food and Drug Administration);
- 3) les autres techniques de séquençage d'ADN ne correspondant pas à la définition du SNG, comme le séquençage de type Sanger. Celui-ci fait l'objet d'un document indépendant publié par l'INESSS en septembre 2019 [INESSS, 2019]

2. MÉTHODOLOGIE

2.1. Question d'évaluation

La question d'évaluation guidant le présent travail est la suivante :

Quelles sont les meilleures pratiques de validation analytique de tests cliniques effectués par séquençage de nouvelle génération (SNG) en génétique humaine?

2.2. Revue de littérature

Une revue de la littérature scientifique et grise concernant la validation analytique des tests SNG a été réalisée. Cette revue de littérature ciblait principalement les publications provenant d'organismes avec comités de rédaction internationaux (ex. : lignes directrices, guides et recommandations de sociétés savantes et réglementaires), les publications d'associations professionnelles ainsi que les documents de synthèse.

Les bases de données MEDLINE (Pubmed), Embase (Ovid) et EBM Reviews (Ovid) ont été interrogées afin de repérer des études de langue anglaise et française publiées entre janvier 2000 et mai 2019. La stratégie de recherche d'information a été élaborée en collaboration avec un spécialiste en information scientifique. Les bibliographies des publications retenues ont été consultées afin de répertorier d'autres documents pertinents. Le détail des différentes stratégies de recherche d'information est présenté à l'[annexe A](#).

Seuls les documents présentant des recommandations en matière de validation analytique d'un test SNG ont été retenus. La sélection a été effectuée par un professionnel scientifique, d'abord par les titres et les résumés, puis sur le texte intégral. Les données ont été extraites par deux professionnels scientifiques de l'INESSS qui se sont séparés les documents à extraire. Aucune évaluation de la qualité méthodologique des documents recensés n'a été réalisée. Une synthèse narrative des données extraites a été réalisée.

2.3. Processus de participation

L'INESSS a mis sur pied un comité consultatif dont le mandat consistait à proposer, aux membres du CSABM, un cadre de référence pour la validation analytique des techniques de diagnostic moléculaire par séquençage d'ADN, dont le SNG. Plus précisément, le comité a revu la littérature sur le sujet, élaboré la liste des éléments devant faire partie de la validation analytique, et défini, pour chaque élément, les paramètres, mesures et informations requis en matière de validation analytique d'une nouvelle analyse de diagnostic moléculaire. L'équipe de l'INESSS a également consulté trois lecteurs externes pour valider le contenu du document et l'applicabilité des recommandations dans le contexte québécois.

Les noms de toutes les personnes consultées sont énumérés dans les pages liminaires du présent rapport. Une déclaration de conflits d'intérêts et de rôles potentiels a été recueillie pour chacun des membres du comité et des lecteurs externes. Aucun n'a rapporté de conflit d'intérêts potentiels susceptibles d'affecter leur rôle. La version finale du document n'engage pas la responsabilité des différentes personnes consultées.

3. PRÉAMBULE

La validation analytique d'un test de laboratoire exige de fournir les éléments de preuve, procurant un haut degré d'assurance, que ce dernier performera toujours comme attendu selon son utilisation prévue [CAP, 2015; CLSI, 2012a].

Une « validation analytique » exhaustive doit être effectuée pour tout nouveau test conçu localement en laboratoire ou TCL (en anglais *Laboratory Developed Test* ou LDT) avant qu'il soit utilisé en contexte diagnostique ou clinique [CSA, 2018; ISO, 2012]. Les situations d'application des TCL et des précisions sont présentées au tableau 1.

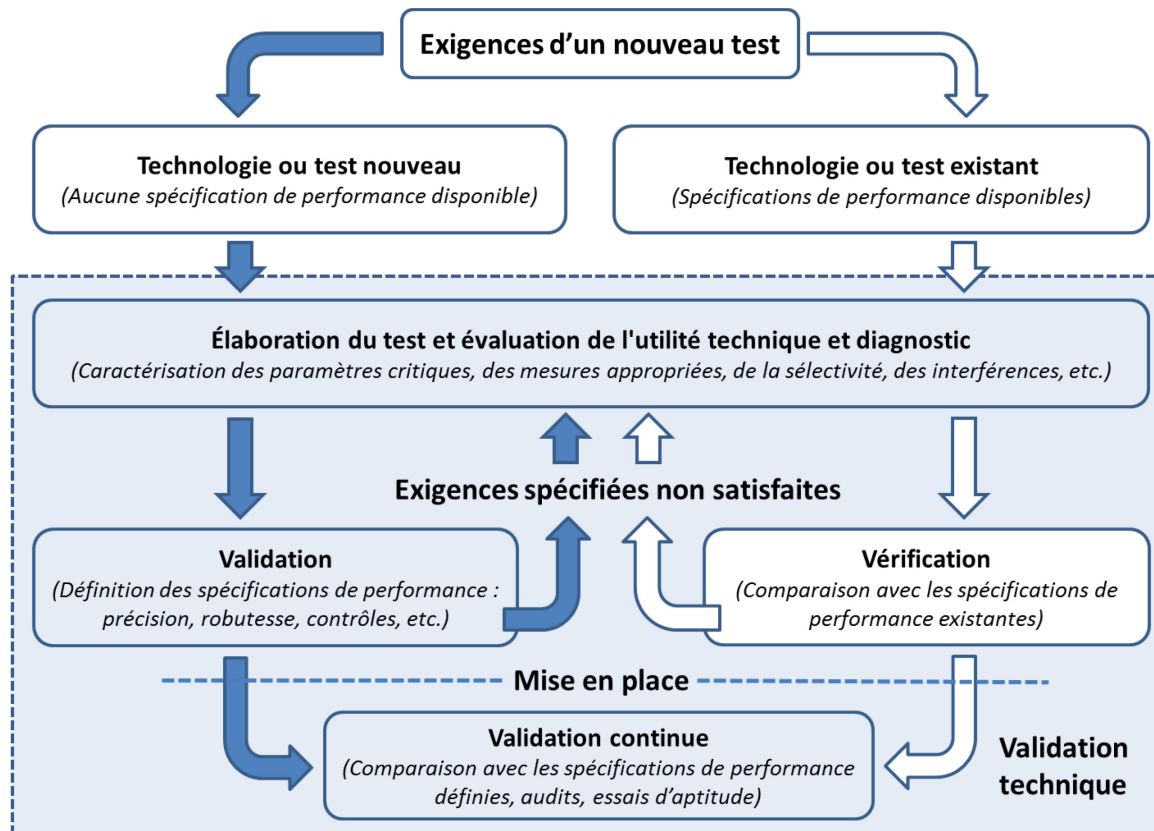
Tableau 1 Tests conçus en laboratoire (TCL)

SITUATIONS D'APPLICATION	PRÉCISIONS
Test élaboré à partir des principes premiers (<i>de novo</i>)	Le laboratoire ou le réseau de laboratoires est responsable de la conception et de la réalisation de l'essai.
Test conçu ou modifié à partir d'un protocole publié ou d'une autre source	Le test est effectué : i) conformément à la littérature scientifique; ii) à partir des spécifications de conception d'un TCL conçu par un autre laboratoire; iii) à partir des spécifications de conception de toute autre source.
Test constitué d'un assemblage de composantes disponibles sur le marché ou de réactifs spécifiques à l'analyte	Sans objet
Test normalisé disponible sur le marché utilisé à une autre fin que celle prévue par le fabricant	Un dispositif de diagnostic <i>in vitro</i> (DDIV) doit être considéré comme un TCL si : i) l'utilisation prévue est différente de celle déclarée par le fabricant; ii) un composant physique du DDIV commercial est modifié, remplacé ou enlevé; iii) l'essai n'est pas utilisé conformément aux directives du fabricant.

Source : CSA, 2018.

Dans le cas d'une analyse ayant déjà fait l'objet d'une validation analytique et dont la performance a été établie (ex. : trousse commerciale homologuée ou analyse déjà validée localement), on ne parle pas de « validation analytique » mais plutôt de « vérification analytique ». La figure 1 présente le processus de mise en place d'un nouveau test diagnostique, y compris les exigences concernant la validation ou la vérification du test en fonction de la disponibilité ou non des spécifications de performance appropriées [Mattocks *et al.*, 2010].

Figure 1 **Processus de mise en place d'un nouveau test de génétique moléculaire aux fins d'une utilisation en diagnostic médical**



Source : Schéma adapté de celui publié par Mattocks et ses collaborateurs [2010].

Bien que les différents organismes réglementaires et professionnels s'entendent sur la nécessité d'effectuer la validation ou la vérification analytique de nouveaux tests de séquençage, les paramètres et exigences minimales requises diffèrent d'un organisme à l'autre.

Les différentes étapes et les principaux critères et paramètres nécessaires à la validation des essais effectués par SNG et abordés dans ce document se réfèrent notamment aux publications des organismes suivants :

Organismes ayant des comités de rédaction internationaux et une reconnaissance internationale :

- le Clinical and Laboratory Standards Institute (CLSI), soit les documents MM09, MM01-A3, MM20-A et EP12-A2 [CLSI, 2014; CLSI, 2012a; CLSI, 2012b; CLSI, 2008];

Organismes réglementaires :

- la Food and Drug Administration (FDA) [FDA, 2018; FDA, 2016];

Associations professionnelles (ou de chercheurs) :

- le Collège canadien des généticiens médicaux (CCGM) [Hume *et al.*, 2019];
- l'International Quality Network for Pathology (IQN Path) [Deans *et al.*, 2017; Deans *et al.*, 2015];
- l'Institut national du cancer (INCa) [INCa, 2016; INCa, 2014];
- l'Association for Molecular Pathology et le College of American Pathologists [Roy *et al.*, 2018; Jennings *et al.*, 2017; Li *et al.*, 2017];
- le groupe de travail Nex-StoCT (de l'anglais *Next-generation Sequencing : Standardization of Clinical Testing*) [Gargis *et al.*, 2012];
- le College of American Pathologists (CAP) [Aziz *et al.*, 2015];
- l'American College of Medical Genetics and Genomics (ACMG) [Rehm *et al.*, 2013];
- l'Association for Clinical Genetic Science (ACGS) [Deans *et al.*, 2015];
- l'EuroGentest Validation Group [Matthijs *et al.*, 2016; Mattocks *et al.*, 2010].

4. OBJECTIF DE PERFORMANCE RECOMMANDÉ PAR LE GROUPE DE TRAVAIL

Le groupe de travail a établi, dès le début des travaux, que l'objectif de performance d'un test de séquençage devrait être de produire une erreur dans la séquence rapportée moins d'une fois sur cent rapports-patients ($< 1\%$ des rapports). En conséquence, pour obtenir un tel résultat, le taux d'erreur à chaque position de la séquence doit être inférieur à 1% divisé par la longueur totale de la séquence analysée [Rousseau *et al.*, 2009]. Davantage d'informations à ce sujet sont disponibles à la [section 7.3](#) de ce document.

5. UTILISATION CLINIQUE ET TYPE D'ANALYSE

5.1. Utilisation clinique du SNG

Compte tenu de sa capacité de séquençage élevée, le SNG peut être utilisé pour analyser un seul gène, un panel de gènes plus ou moins important, l'exome ou le génome entier [Aziz *et al.*, 2015; ANPGM, 2014; Gargis *et al.*, 2012]. Une description du principe technologique du SNG est présentée à l'[annexe B](#).

Dans le domaine du diagnostic moléculaire des maladies héréditaires, les techniques de SNG sont notamment utilisées pour préciser la cause génétique d'une pathologie présentant une certaine hétérogénéité ou pour mettre fin à l'errance diagnostique [Deans *et al.*, 2015; ANPGM, 2014]. Le séquençage de grands gènes ou de multiples gènes (panel de gènes) est l'application la plus répandue du SNG en diagnostic clinique. Cette approche vise à détecter des variations pathogéniques dans les gènes connus comme étant associés à la condition du patient, tout en réduisant au minimum les possibilités de détecter des variations nucléotidiques dans des gènes à pénétrance élevée sans lien avec le phénotype clinique du patient, ou dans ceux dont la signification clinique n'a pas encore été établie [Deans *et al.*, 2015].

Afin de rationaliser et d'optimiser le travail en laboratoire, ou d'assurer une utilisation efficace de certains échantillons (ex. : biopsie invasive, matériel peu abondant, rare ou unique), le séquençage d'un plus grand nombre de gènes est parfois effectué, suivi par l'analyse ciblée de gènes ou de régions d'intérêt. Notamment, l'utilisation de panels contenant plusieurs gènes de prédisposition aux cancers héréditaires ainsi que le séquençage complet de l'exome sont des stratégies de capture utilisées pour l'analyse ultérieure de sous-panels de gènes additionnels [Deans *et al.*, 2015].

Outre le diagnostic des maladies héréditaires, le SNG est aussi utilisé dans les domaines de l'oncologie (ex. : tumeurs solides, hémopathies malignes comme les leucémies), de la compatibilité du système HLA (de l'anglais *human leukocyte antigen*), des maladies infectieuses et du dépistage prénatal non invasif pour détecter les anomalies chromosomiques du fœtus [Aziz *et al.*, 2015].

Le SNG permet la détection simultanée d'une variété d'aberrations génomiques telles que :

- les variations d'un seul nucléotide (VSN ou SNV, de l'anglais *single nucleotide variant*);
- les variations de plusieurs nucléotides (VPN ou MNV, de l'anglais *multiple nucleotides variants*);
- les grandes insertions ou délétions;
- certains réarrangements structuraux (translocations, inversions);
- certains homopolymères et certaines séquences répétitives;

- les variations du nombre de copies (VNC ou CNV, de l'anglais *copy number variants*) [Aziz *et al.*, 2015; Luthra *et al.*, 2015; Salto-Tellez et Gonzalez de Castro, 2014].

5.2. Type d'analyse

La recherche d'une variation pathogénique par séquençage est généralement considérée comme une méthode qualitative, puisqu'à chaque position de la séquence analysée, seulement quatre catégories de résultat sont possibles, soit les quatre bases nucléotidiques de l'ADN : l'adénine (A), la thymine (T), la cytosine (C) et la guanine (G). Il s'agit de détecter la présence d'une variation qualitative de séquence nucléotidique par rapport à une séquence de référence [INCa, 2014; Mattocks *et al.*, 2010]. Ainsi, un nucléotide concordant avec la séquence de référence humaine est considéré comme « normal » et les trois autres nucléotides possibles sont considérés comme des variations nucléotidiques [Pont-Kingdon *et al.*, 2012].

Étant donné la nature diploïde du génome humain, trois statuts différents peuvent être déterminés pour toute variation constitutionnelle de séquence, soit « homozygote normal », « variation nucléotidique hétérozygote » ou « variation nucléotidique homozygote ». La détermination du caractère homozygote, hémizygote (présence d'un seul allèle au lieu de deux) ou hétérozygote d'une variation nucléotidique se fait en fonction de la détection ou non de l'allèle normal et de l'allèle variant, et ce, en considérant le mode de transmission. Le tableau 2 présente les différents résultats et profils génétiques pouvant être obtenus par séquençage.

Tableau 2 Résultats et profils génétiques pouvant être obtenus à la suite de l'analyse de l'ADN par séquençage

NUCLÉOTIDE DE RÉFÉRENCE	VARIATION NUCLÉOTIDIQUE	RÉSULTAT (PROFIL)
Présent (+)	Absente (-)	Négatif (homozygote normal ou hémizygote normal)
Présent (+)	Présente (+)	Positif (hétérozygote)
Absent (-)	Présente (+)	Positif (homozygote muté ou hémizygote muté)

Source : Tableau inspiré de celui publié par l'Association des praticiens en génétique moléculaire [ANPGM, 2012].

Précisons toutefois que le SNG peut également produire des informations de type quantitatif, notamment lors de la détection de variations du nombre de copies ou de ratios chromosomiques en dépistage prénatal non invasif.

De plus, des critères quantitatifs de performance servant à évaluer la qualité d'exécution des procédés analytiques, des instruments et des processus informatiques, ainsi que la qualité des données de séquençage produites, doivent être déterminés lors des études de validation appropriées [Aziz *et al.*, 2015; Gargis *et al.*, 2012]. Ces critères de qualité sont présentés à la [section 7](#) de ce rapport.

6. PLANIFICATION PRÉANALYTIQUE

Selon les normes ISO 15189 et CAN/CSA Z316.8-18 [CSA, 2018; ISO, 2012], toute validation analytique d'un test nécessite une planification préanalytique. Cette dernière comporte les étapes suivantes :

- 1) Devis et conception du test. Cette étape consiste à :
 - définir ce que la méthode doit réaliser;
 - spécifier les éléments de la méthode, soit :
 - la validité scientifique du test (validité et utilité cliniques);
 - l'usage prévu du test (population, critères cliniques);
 - les méthodes de comparaison (méthode de référence, étude de concordance avec les résultats obtenus d'une autre méthode ou par un autre laboratoire).
- 2) Élaboration du plan directeur de validation (de l'anglais *master validation plan*). Le plan directeur de validation doit définir les objectifs de la méthode et les critères de performance à mesurer lors de la validation analytique, ainsi que les critères de qualité prédéterminés pour chacune des dimensions de la performance du test qui seront mesurées.
- 3) Données sur la qualité (cinq étapes essentielles à réaliser) :
 - Qualification des instruments analytiques;
 - Validation des méthodes analytiques;
 - Tests de conformité des systèmes;
 - Vérifications et contrôles de qualité (ex. : qualité des réactifs, certificats d'analyse);
 - Évaluation du risque à chaque étape d'exécution du test.

En particulier, pour une analyse effectuée par SNG, les éléments suivants sont également essentiels :

- Détermination des séquences de référence;
- Conception, choix et validation des amorces (dans le cas d'un TCL);
- Détermination du matériel de référence et des échantillons cliniques témoins.

Ces différentes étapes sont résumées dans les sections suivantes.

6.1. Détermination des séquences de référence

La séquence de référence doit être la plus récente possible et devrait être obtenue du National Center for Biotechnology Information (NCBI)¹ ou du répertoire Ensembl², et idéalement provenir du Locus Reference Genomic (LRG)³ [Ellard *et al.*, 2016]. Si aucune séquence du LRG n'est disponible, l'ACGS recommande d'utiliser les séquences RefSeq⁴ parce qu'elles seraient plus complètes et mieux annotées que d'autres séquences GenBank^{MC}, en plus de ne pas être redondantes (une seule séquence par gène ou isoforme) [Ellard *et al.*, 2016]. La possibilité qu'il existe d'autres transcrits pertinents, sur le plan biologique, ou des pseudogènes, doit être prise en considération [Ellard *et al.*, 2016; Pont-Kingdon *et al.*, 2012]. Il est recommandé que les coordonnées génomiques d'une séquence de référence, y compris le numéro de construction du génome assemblé ainsi que sa version, soient enregistrées (ex. : GRCh38.p9).

Dans un deuxième temps, les gènes ou les régions du génome à analyser doivent être déterminés et circonscrits [Luthra *et al.*, 2015]. La description de ces régions, incluant les variations pathogéniques ou les points chauds (de l'anglais *hot spot*) qu'elles comportent, permet de définir les intervalles de référence (*reference range*) et de mesure (*reportable range*) de l'analyse moléculaire [Pont-Kingdon *et al.*, 2012].

Lorsque des données concernant la distribution des variations pathogéniques ne sont pas disponibles, le CAP recommande que tous les exons, incluant 20 à 50 nucléotides des séquences introniques flanquantes, soient analysés [Pont-Kingdon *et al.*, 2012]. Pour sa part, l'ACGS indique que la couverture minimale inclut les régions codantes du gène (exons), les sites d'épissage conservés, ainsi que les 10 premiers nucléotides des séquences introniques situées en amont et en aval de chacun des exons [Ellard *et al.*, 2016]. Le séquençage des régions promotrices, des régions non traduites en 3' (telles que les sites de polyadénylation) et les régions introniques peut présenter un intérêt si les variations pathogéniques contenues dans ces régions sont bien caractérisées [Ellard *et al.*, 2016; Pont-Kingdon *et al.*, 2012]. Précisons que l'assignation d'une probabilité de pathogénicité aux variations nucléotidiques introniques éloignées de la frontière intron-exon est difficile et incertaine. Les variations plus éloignées que 10 nucléotides en amont et en aval des exons ne sont généralement pas rapportées⁵.

¹ National Center for Biotechnology Information (NCBI) [site Web], disponible à : <https://www.ncbi.nlm.nih.gov/>.

² Ensembl GRCh37. Human [site Web], disponible à : https://grch37.ensembl.org/Homo_sapiens/Info/Index.

³ Les séquences du RefSeq (Reference Sequence) sont des copies de génomes, disponibles sur GeneBank^{MC}, qui sont sélectionnées et assemblées. Les transcrits et les protéines RefSeq sont générés par des méthodes qui comprennent des calculs informatiques, une édition manuelle du contenu et la propagation de génomes annotés soumis aux membres de l'International Nucleotide Sequence Database Collaboration (INSDC). Locus Reference Genomic (LRG) [site Web], disponible à : <http://www.lrg-sequence.org/>.

⁴ RefSeq: NCBI Reference Sequence Database [site Web], disponible à : <https://www.ncbi.nlm.nih.gov/refseq/>.

⁵ Information fournie par un des experts consultés.

L'étendue de l'analyse doit donc être définie. Une région minimale de couverture horizontale doit inclure les régions codantes du gène et les sites d'épissage [Deans *et al.*, 2015]. Selon l'ACGS, il est plus simple pour le laboratoire et ceux qui recevront les résultats d'utiliser une distance limite standard « intron-exon » pour tous les gènes d'un même panel. Si cette région doit être prolongée ou réduite pour une raison quelconque, des justifications devraient être fournies (ex. : limitations techniques, obligation de couvrir une variation pathogénique connue, etc.) [Deans *et al.*, 2015].

Précisons que certaines régions du génome humain sont difficiles, voire impossibles à séquencer par SNG [Deans *et al.*, 2015], notamment:

- les gènes pour lesquels la présence de séquences hautement homologues (ex. : pseudogènes) complexifie ou rend impossible une analyse ciblée;
- les régions riches en GC (ex. : les régions promotrices et le premier exon des gènes sont particulièrement riches en CG);
- les éléments répétitifs.

6.2. Conception, validation et choix des amorces

Les laboratoires offrant des services de SNG peuvent le faire à partir de panels commerciaux offerts par le fournisseur de leur plateforme ou à partir de panels conçus sur mesure selon leurs besoins et exigences [Luthra *et al.*, 2015].

Les deux approches comportent des avantages et des inconvénients. Les arguments en faveur des panels commerciaux, par rapport aux panels sur mesure, sont leur disponibilité immédiate, leur validation analytique (dans le cas d'une trousse normalisée) et leur optimisation préalable, ainsi que la communauté d'utilisateurs plus large, ce qui permet une plus grande standardisation des résultats. Cependant, les dispositifs de diagnostic *in vitro* (DDIV) normalisés exigent malgré tout une vérification analytique par le laboratoire qui souhaite les utiliser. Enfin, la correspondance entre les panels commerciaux et les besoins locaux peut être sous-optimale [Van den Bulcke *et al.*, 2015].

Le laboratoire qui désire développer un panel sur mesure (donc un TCL) doit s'assurer que les amorces utilisées permettent de couvrir l'intégralité des régions d'intérêt et qu'il minimise les biais de capture et le risque de perte allélique. Une évaluation préliminaire *in silico* peut être effectuée à l'aide de la séquence des amorces que l'on propose d'utiliser. Par la suite, en cours de validation et lors de chaque essai, une évaluation du biais associé aux brins sens et antisens de l'ADN doit être effectuée ([tableau C-2](#), annexe C). En effet, seule la mesure empirique du taux de perte allélique spécifique à chaque locus amplifié permet d'avoir une estimation fiable de ce risque. Moins il y a d'amplification des cibles préalablement au séquençage, plus le risque de perte allélique est faible, car la perte allélique est propre aux techniques d'amplification des acides nucléiques (TAAN).

Les critères de sélection des amorces sont généralement les mêmes que ceux utilisés en PCR classique et en séquençage Sanger [INCa, 2016]. Il est notamment recommandé d'éviter les sites de polymorphisme avec une fréquence de plus de 1 % dans la population, et ce, dans les 3 derniers nucléotides de l'amorce. Selon l'Association

nationale des praticiens de génétique moléculaire (ANPGM), une tolérance maximale de 1 % de polymorphismes sur les 5 derniers nucléotides de l'amorce serait recommandée pour les activités en génétique constitutionnelle. Une vérification à cet égard doit être reconduite à intervalles réguliers (tous les 2 ans), car les bases de données évoluent régulièrement [ANPGM, 2014].

Pour les techniques de capture, les zones couvertes par les amorces doivent être vérifiées. Puisqu'environ 92 % des variations pathogéniques introniques se trouvent à l'intérieur des 10 premières paires de bases (pb) de part et d'autre de l'exon, l'ACGS recommande de considérer l'analyse de ces dernières en plus des exons codants comme région minimale interrogée lors de la conception de nouveaux essais de capture [Deans *et al.*, 2015]. Un contrôle permanent de la couverture et de la profondeur de lecture sur les cibles doit être documenté [INCa, 2016].

6.3. Détermination des types de spécimens cliniques acceptés pour un test effectué par SNG

Le choix des spécimens cliniques utilisés pour le séquençage doit être guidé par le type de variations nucléotidiques recherché, soit les variations constitutives ou somatiques. Comme le séquençage présente certaines limites techniques quant à la détection de variations en mosaïque ou clonales, la sélection et le traitement approprié de la source d'ADN sont essentiels à la validité de l'analyse et à son interprétation.

Le prélèvement sanguin comme source de cellules leucocytaires constitue le matériel biologique typiquement utilisé pour la recherche de variations nucléotidiques constitutives [CLSI, 2014]. Pour la recherche de variations nucléotidiques somatiques ou en mosaïque, d'autres types de spécimens peuvent être utilisés, tels qu'un prélèvement de tissu tumoral ou de moelle osseuse [CLSI, 2012b].

Précisons qu'en raison de la réticulation (de l'anglais *crosslinking*) de l'ADN causée par le formaldéhyde, le séquençage effectué à partir d'ADN extrait de tissus fixés dans la formaline et enrobés de paraffine (ou FFEP) est plus sensible aux artefacts qui peuvent potentiellement être la cause de résultats faussement positifs. Ce type d'échantillon est souvent utilisé dans le domaine de l'oncologie [Ellard *et al.*, 2016; CAP, 2015].

6.4. Détermination du matériel de référence et des échantillons cliniques témoins nécessaires à la validation analytique d'un test effectué par SNG

Selon le CLSI, la validation analytique d'une méthode ou d'un test en génétique moléculaire doit être effectuée à l'aide de matériel de référence et d'échantillons cliniques témoins [CLSI, 2012a].

Le matériel de référence peut être constitué d'échantillons issus d'une population dite « normale » et d'une population qui présente les variations nucléotidiques pertinentes. Concernant cette dernière, le matériel de référence devrait contenir les variations

nucléotidiques de la population d'intérêt ou observées dans un contexte similaire à celui de la population ciblée [CLSI, 2014].

Des échantillons cliniques témoins doivent être inclus à titre de témoins internes positifs et négatifs lors de la validation analytique du test et dans chaque analyse de routine pour détecter les erreurs possibles de préparation des échantillons ainsi que les échecs de réaction d'amplification ou de séquençage. Les échantillons cliniques utilisés comme témoins doivent être du même type que ceux qui seront analysés dans le cadre clinique et la méthode d'extraction de l'ADN doit, dans la mesure du possible, être la même que celle utilisée par le laboratoire qui effectuera le test [Pont-Kingdon *et al.*, 2012].

Le nombre, le type et la variété des échantillons cliniques utilisés lors de la validation doivent être adéquats, de manière à s'assurer que les résultats pourront être interprétés selon les caractéristiques de la population ciblée et que les limites du test sont connues (le nombre d'échantillons souhaitable est abordé dans la [section 7.5](#) de ce document).

Le tableau 3 présente quelques recommandations du CLSI concernant le matériel de référence et les échantillons cliniques témoins à utiliser lors de la validation analytique d'un test de génétique moléculaire conçu en laboratoire [CLSI, 2012a].

Tableau 3 Recommandations du CLSI concernant le matériel de référence et les échantillons cliniques à utiliser lors de la validation analytique d'un test de génétique moléculaire

ÉCHANTILLONS	RECOMMANDATIONS DU CLSI
Matériel de référence	Considérer la provenance du matériel de référence selon l'ordre d'importance décroissant suivant : 1. Matériel de référence certifié, lorsque celui-ci est disponible; 2. Échantillons testés par un laboratoire de référence accrédité; 3. Échantillons testés par un laboratoire sous-traitant accrédité; 4. Échantillons testés par une méthode reconnue et fiable.
Échantillons cliniques témoins	▪ Nombre adéquat ▪ Type représentatif des spécimens cliniques qui seront analysés (ex. : sang, frottis buccal, tissu frais ou congelé, tissus fixés et enrobés de paraffine, cellules en culture, etc.) ▪ Variété représentative de la population ciblée (ex. : sexe, âge, ethnicité, condition clinique) ▪ Variété représentative de tous les résultats possibles (ex. : VSN, VMN, VNC, insertion, délétion, duplication, translocation)

Abréviations : CLSI : Clinical and Laboratory Standard Institute; VNC : variation du nombre de copies; VMN : variation de multiples nucléotides; VSN : variant d'un seul nucléotide.

Dans le contexte de la validation analytique de tests effectués par SNG, l'utilisation de références commerciales devrait être envisagée, notamment pour assurer la détection de variations nucléotidiques rares ou complexes. Comparativement aux échantillons cliniques, qui sont souvent disponibles en quantité limitée, le matériel synthétique ou les lignées cellulaires commerciales de référence procurent de grandes quantités d'acides

nucléiques qui peuvent servir à l'établissement de plusieurs paramètres de performance en plus de servir de contrôle de qualité interne lors des essais [Gargis *et al.*, 2012].

Entre autres, le Genetic Testing Reference Material Coordination Program, le National Institut of Health (NIH) et le Genome in a Bottle Consortium du National Institute of Standards and Technology (NIST) offrent plusieurs types de matériels de référence (ex. : échantillons, fichiers électroniques de données brutes) permettant de valider et d'évaluer de multiples variations nucléotidiques uniques du génome humain. Ces références permettent notamment d'assurer une meilleure représentativité des différents types de variations nucléotidiques qui pourraient être rencontrés dans les spécimens cliniques.

Selon le groupe de travail Nex-StoCT (*Next-generation Sequencing: Standardization of Clinical Testing*), au moins un échantillon de référence commercial, contenant tous les types de variations nucléotidiques que le test est censé détecter dans les régions interrogées, devrait être évalué lors du processus de la validation analytique d'un nouveau test et comme contrôle de qualité interne des essais [Gargis *et al.*, 2012]. La validation doit évidemment être complétée par l'analyse d'un nombre suffisant d'échantillons cliniques témoins.

7. VALIDATION ANALYTIQUE

La validation analytique peut être divisée en trois grandes étapes [CSA, 2018] :

1. Essais préliminaires

Ces essais devraient permettre d'acquérir l'expérience technique nécessaire pour réaliser le test. Ils devraient également permettre de déterminer les paramètres critiques pouvant affecter la performance du test ainsi que les limitations et les mesures de contrôle nécessaires à considérer.

2. Évaluation de la performance

La performance d'une méthode de SNG ou d'un test doit être évaluée selon les dimensions mentionnées à l'[annexe D](#) de ce document. Celles-ci incluent, entre autres, la sensibilité, la spécificité, la précision, la limite de détection et de quantification, l'incertitude de mesure, l'intervalle de mesure et l'intervalle de référence.

3. Vérification ou validation des logiciels

Si un ou des logiciels sont utilisés à quelque moment au cours des processus préanalytique, analytique et postanalytique, ils doivent faire l'objet d'une **vérification** analytique s'ils sont déjà homologués par Santé Canada. Les logiciels commercialisés portant la désignation « *recherche uniquement* », les logiciels conçus par le laboratoire ou les logiciels sous licence, mais utilisés pour un usage différent de celui indiqué par le fabricant, devraient faire l'objet d'une **validation** analytique par le laboratoire, entre autres de leurs fonctionnalités et de leur stabilité.

7.1. Particularités du SNG

En séquençage, les données de validation peuvent être utilisées pour évaluer la performance de la technologie en général (ex. : validation de la méthode ou de la plateforme de séquençage, incluant les systèmes informatiques) ou pour un test en particulier (ex. : validation de la capacité d'un test à détecter des variations pathogéniques dans les gènes *BRCA1* et *BRCA2*) [Mattocks *et al.*, 2010]. Selon le CLSI, la validation de la méthode ne suffit pas à démontrer que le séquençage est capable de détecter toutes les variations nucléotidiques associées à une maladie en particulier. Une validation adaptée à l'application pour laquelle le séquençage sera utilisé doit être également entreprise. Dans le cas du SNG, cela nécessite généralement une démonstration que les variations nucléotidiques qui sont pertinentes sur le plan clinique peuvent être détectées à l'aide du nouveau test [CLSI, 2014].

Toutefois, il a été proposé d'utiliser une approche de validation globale approfondie de la méthode, complétée par des expériences de validation additionnelles moins exhaustives, spécifiques aux différents gènes, exons ou marqueurs à étudier, et de s'assurer que les critères de qualité, concernant notamment la qualité des données de séquence produites par un nouveau test, sont satisfaits [CLSI, 2014; INCa, 2014; Mattocks *et al.*, 2010].

Cette façon de faire, proposée notamment par Mattocks et ses collaborateurs [2010], s'appuie sur le fait qu'en général, la méthode de validation implique d'évaluer la performance du test en comparaison avec une méthode de référence capable de déterminer, hors de tous doutes, le statut positif ou négatif de l'échantillon. Toutefois, le séquençage diffère de la plupart des autres types d'analyses du fait qu'il cible et investigue une multitude d'analytes, soit des milliers de nucléotides individuels lors d'un même essai [CAP, 2015]. De ce fait, la sensibilité et la spécificité d'un test de séquençage sont difficiles à déterminer de façon rigoureuse en raison de l'absence de témoins positifs pour chaque base potentiellement mutée [INCa, 2014].

De plus, l'accumulation de données statistiquement significatives n'est pas toujours possible, notamment dans le cas de l'analyse de maladies rares [Mattocks *et al.*, 2010; Jennings *et al.*, 2009; Maddalena *et al.*, 2005]. Pour ces raisons, Mattocks et ses collaborateurs [2010] recommandent que la validation d'un nouveau test soit plutôt abordée comme une vérification, c'est-à-dire s'assurer que les processus et les données de séquençage satisfont aux normes et critères de qualité établis. Selon les auteurs, les facteurs à considérer devraient inclure :

- la confirmation que le nouveau test cible spécifiquement la région d'intérêt (ex. : vérifier la spécificité des amorces pour les séquences ciblées);
- la confirmation que les deux allèles sont amplifiés de façon fiable (ex. : veiller à ce qu'aucun polymorphisme ne soit situé dans les sites de liaison des amorces);
- la confirmation que les données de séquençage générées sont de bonne qualité sur l'ensemble des régions interrogées (ex. : contrôle de l'atteinte des seuils de qualité minimale établis, qui sera abordée à la [section 7.2](#)).

De ce fait, le processus global de validation analytique des activités de SNG d'un laboratoire clinique est généralement divisé en trois étapes [Gargis *et al.*, 2012] :

- 1) Validation de la méthode de séquençage;
- 2) Validation des processus informatiques (pipeline informatique);
- 3) Validation d'un test spécifique.

Chacune de ces étapes est brièvement décrite dans les sous-sections suivantes. Les paramètres de qualité du séquençage ainsi que certains critères de performance à évaluer lors de la validation analytique de la méthode de SNG ou d'un test en particulier sont respectivement détaillés dans les sections [7.2](#) et [7.4](#) de ce rapport.

7.1.1. Validation de la méthode de SNG

L'objectif de la validation de la méthode de SNG est de caractériser et d'établir les performances de cette dernière et de démontrer, par des mesures empiriques, que la méthode performe selon (ou surpasse) les qualifications minimales requises pour une application proposée [CLSI, 2014].

Le laboratoire doit réaliser la validation de sa méthode de SNG en fonction des variations nucléotidiques représentatives de son activité. À cette étape, les conséquences d'une variation nucléotidiques à l'égard de sa pathogénicité ne sont pas pertinentes. Il est donc approprié d'inclure des variations bénignes ainsi que des variations pathogéniques, pourvu que celles-ci couvrent tous les types de variation de séquence que le laboratoire propose d'identifier à l'aide de sa plateforme de SNG [Deans *et al.*, 2015; Rehm *et al.*, 2013]. Ces variations nucléotidiques doivent idéalement être réparties sur les différents gènes ou exons ciblés par le test [INCa, 2016]. Les références commerciales ou les échantillons certifiés permettent généralement de satisfaire ces exigences.

Pendant cette étape, les paramètres de performance des instruments et du processus d'analyse (c.-à-d. les valeurs utilisées pour juger de la qualité des données et de la fiabilité des résultats de séquençage) doivent être établis pour une application spécifique [CLSI, 2014]. Chacune des étapes de la méthode de SNG doit être validée [Gargis *et al.*, 2012], notamment :

- l'extraction de l'ADN (ou de l'ARN) et la préparation de la librairie;
- la méthode de capture ou d'enrichissement des séquences, s'il y a lieu;
- le processus de séquençage;
- l'analyse des données.

7.1.2. Validation du pipeline informatique

Cette étape permet d'établir et de documenter les paramètres informatiques qui sont appropriés pour la génération de données de séquence fiables et de qualité. Elle permet également de déterminer la capacité des logiciels à identifier et à classer les variations de séquence détectées [Gargis *et al.*, 2012].

Le pipeline informatique nécessaire au traitement et à l'analyse des données de SNG incorpore de multiples outils. Il comprend généralement des logiciels destinés à :

- l'appel des bases;
- l'alignement et l'assemblage des séquences;
- l'appel des variations nucléotidiques;
- l'annotation des séquences;
- l'identification des variations de séquence cliniquement pertinentes [Matthijs *et al.*, 2016; CLSI, 2014].

Le groupe de travail Nex-StoCT recommande que le pipeline informatique d'une plateforme soit optimisé en tant qu'entité distincte lors de la validation de la plateforme de séquençage et lors du développement d'un test [Gargis *et al.*, 2012]. Afin de valider spécifiquement l'étape de l'analyse des données de séquence, des échanges de données électroniques entre laboratoires ou l'obtention de fichiers de données de référence commerciales peuvent être envisagés [Deans *et al.*, 2015; Gargis *et al.*, 2012].

7.1.3. Validation d'un test

La validation d'un test en particulier consiste à mesurer la performance analytique de ce dernier et à démontrer qu'il satisfait les critères de performance prédéfinis dans le plan directeur de validation de la méthode et du pipeline informatique [Deans *et al.*, 2017]. L'objectif est de montrer que la performance du test est adéquate pour ses indications d'utilisation. Cela implique généralement d'évaluer si le test identifie ou mesure, dans les limites statistiques définies, la présence ou l'absence de variations de séquence dans les régions ciblées du génome [FDA, 2018].

Il est essentiel que les témoins positifs utilisés lors de la validation du test comportent une gamme de variations pathogéniques représentatives de la pathologie à l'étude. Comme il a été mentionné précédemment, cette étape peut être effectuée par l'utilisation de matériel de référence commercial. Toutefois, des échantillons biologiques de patients, du même type que ceux qui seront analysés, doivent également être utilisés dans le processus de validation d'un test. Idéalement, ces échantillons biologiques devraient comporter, dans la mesure du possible, des variations pathogéniques connues comme étant associées à la pathologie à l'étude [Deans *et al.*, 2015]. Les études de validation devraient être menées à l'aveugle à l'égard des variations nucléotidiques testées [Deans *et al.*, 2017].

7.2. Évaluation de la qualité des données obtenues par SNG

Une variété de paramètres et de mesures servent à évaluer la qualité des résultats obtenus par SNG [Deans *et al.*, 2015]. Une combinaison de ces paramètres doit être étudiée et des seuils de qualité minimaux doivent être établis dans le plan directeur de validation de la méthode et du pipeline informatique et mesurés lors de la validation analytique afin de garantir la performance des processus. Ces seuils ou mesures de qualité devront être atteints (ou surpassés) au moment d'effectuer chaque essai et servent donc de barème afin de discriminer les données de bonne et de mauvaise qualité [Matthijs *et al.*, 2016; CLSI, 2014].

Les tableaux [C-1](#) et [C-2](#) de l'annexe C résument certains des principaux paramètres permettant d'assurer la qualité des données de SNG. Précisons que cette liste n'est pas exhaustive et que plusieurs autres paramètres existent, parfois spécifiques à la chimie et à la méthodologie de SNG utilisée. Le tableau 4 reprend certains de ces critères jugés nécessaires, par le groupe de travail Nex-StoCT, à l'évaluation de la performance d'un essai de SNG [Jennings *et al.*, 2017; Gargis *et al.*, 2012]. Les considérations concernant la validation analytique de ces paramètres et leur utilisation comme contrôles de qualité continue y sont également décrites.

Deux paramètres généralement jugés nécessaires à l'appel d'une variation nucléotidique détectée par un test, soit la valeur de qualité (VQ) de l'appel des bases et la couverture (verticale et horizontale), sont plus amplement décrits dans les paragraphes suivants. Des recommandations d'organismes professionnels et réglementaires concernant les valeurs de qualité minimales de l'appel des bases et de la couverture pour l'appel d'une variation nucléotidique sont présentées au [tableau 6](#).

Tableau 4 Critères nécessaires à l'évaluation des performances analytiques d'un essai de SNG

PARAMÈTRES DE PERFORMANCE*	CONSIDÉRATIONS POUR LA VALIDATION	CONSIDÉRATIONS POUR LES CONTRÔLES DE QUALITÉ CONTINUUS (maintien de l'aptitude à l'emploi)
Couverture (couverture horizontale)	La couverture d'une région particulière, sous des conditions normales d'essai (seuil de couverture), devrait être établie. Il est essentiel qu'une couverture adéquate soit déterminée pour obtenir une sensibilité et une spécificité adéquates pour les régions d'intérêt.	Les résultats rapportés devraient se limiter à ceux générés à l'intérieur de l'intervalle validé du test.
Uniformité de la couverture	La couverture des régions ciblées devant être atteinte pour produire des résultats de séquençage fiables devrait être établie.	L'uniformité de couverture devrait être surveillée et comparée à celle établie lors de la validation. Un profil d'uniformité de couverture qui n'est pas comparable à celui établi lors de la validation pourrait être un indicateur d'erreurs dans le processus d'analyse.
Biais associé aux régions riches en GC	Le contenu en guanine et en cytosine (GC) affecte l'efficacité des réactions de séquençage et affectera la qualité de l'appel des bases et l'uniformité de couverture de la région ciblée. Le biais attribuable au contenu en GC de toutes les régions du génome incluses dans l'analyse devrait être déterminé lors de la validation.	Les biais associés au contenu en GC devraient être surveillés lors de chaque essai afin de détecter les changements de performance du test.
Ratio transitions/transversions	Le ratio des transitions et des transversions (Ti/Tv) devrait être comparable aux valeurs publiées.	Le ratio Ti/Tv devrait être surveillé pour chaque échantillon afin de détecter un changement dans les performances du test. Un ratio Ti/Tv plus faible ou plus élevé que celui attendu est un indicateur d'une faible qualité de l'appel des bases et du fait qu'elle contient potentiellement des erreurs.
Valeur de qualité (VQ) de l'appel des bases	Un seuil de qualité de l'appel des bases acceptable devrait être établi lors de la validation. Des filtres informatiques devraient être mis en place afin de permettre d'éliminer toute lecture comportant des appels de base plus faible que la valeur de qualité établie. Pour les technologies comportant des lectures plus longues, lorsque la détection de larges indels est de mise, les alignements peuvent tolérer de plus faibles valeurs de qualité de l'appel des bases étant donné que la détection d'indels est moins affectée par l'exactitude de chacune des bases sur de longues séquences.	La qualité du signal par rapport au bruit de fond (de l'anglais <i>signal to noise ratio</i> ou SNR) devrait être surveillée en examinant les valeurs de qualité de l'appel des bases et la qualité du SNR d'une lecture pour chaque essai. Note : Les scores de qualité sont difficilement comparables entre les différents instruments ou méthodes de SNG.

PARAMÈTRES DE PERFORMANCE*	CONSIDÉRATIONS POUR LA VALIDATION	CONSIDÉRATIONS POUR LES CONTRÔLES DE QUALITÉ CONTINUS (maintien de l'aptitude à l'emploi)
Qualité de l'alignement	La qualité de l'alignement est une mesure de l'incertitude qu'une lecture soit correctement alignée sur sa position génomique. La démonstration que le test analyse uniquement les lectures s'alignant exclusivement sur les régions spécifiques ciblées par le test devrait être effectuée lors de la validation. Des filtres informatiques devraient être mis en place afin de permettre d'éliminer toute lecture s'alignant sur des régions non ciblées et d'éliminer les duplicatas de lectures.	La proportion des lectures qui ne s'alignent pas sur les régions ciblées devrait être surveillée à chaque essai. Une lecture qui ne correspond pas à la séquence de référence indique que l'échantillon ne performe pas selon les paramètres établis et que cette lecture devrait être exclue de l'analyse. Pour les applications comportant des étapes d'enrichissement, une faible qualité des alignements pourrait être le résultat d'une amplification ou d'une capture non spécifique de l'ADN, ou d'une contamination.
Retrait des duplications des lectures	Des filtres informatiques devraient être mis en place afin de permettre d'éliminer les duplicatas de lectures provenant d'une amplification clonale au moment de l'alignement (à l'exception de celui ayant le score de qualité le plus élevé).	Le retrait des duplications devrait être surveillé afin de prévenir une distorsion des fractions alléliques.
Déclin de l'intensité du signal	Lors de la validation d'un test, l'intensité attendue du signal au cours de la lecture devrait être évaluée afin d'établir les critères de performance normale ainsi que de documenter la diminution attendue de l'intensité du signal.	La diminution attendue de l'intensité du signal devrait être surveillée à chaque essai. Une réduction ou une augmentation soudaine de l'intensité du signal indique une erreur dans la chimie de séquençage.

Sources : Jennings *et al.*, 2017; Gargis *et al.*, 2012.

Abréviations : ADN : acide désoxyribonucléique; GC : guanine et cytosine; SNG : séquençage de nouvelle génération.

* Cette liste n'est pas exhaustive et peut varier en fonction de la plateforme de SNG. De plus, les paramètres et mesures de qualité sont calculés selon des formules et des données qui diffèrent en fonction de la chimie ou de la technologie de séquençage, de la séquence interrogée, du contexte et de la sensibilité d'un test donné, et ne sont généralement pas directement comparables d'une plateforme ou d'un test à l'autre [Deans *et al.*, 2015; CLSI, 2014; Gargis *et al.*, 2012].

7.2.1. Valeur de qualité de l'appel des bases

Dans une analyse impliquant le séquençage, les faux positifs (détection d'une variation pathogénique qui n'existe pas) et les faux négatifs (défaut de détection d'une variation pathogénique présente) peuvent être la conséquence de la qualité insuffisante des profils de séquençage [Mattocks *et al.*, 2010].

Comme il a été mentionné précédemment, la sensibilité et la spécificité d'un test de séquençage sont difficiles à déterminer de façon rigoureuse en raison de l'absence de témoins positifs pour chaque base séquencée. Ces paramètres peuvent cependant être estimés par la valeur de qualité (VQ) obtenue pour chacune des bases de la séquence [INCa, 2014].

La VQ permet d'évaluer la probabilité que l'assignation de la base individuelle soit juste. Elle est exprimée en $\log_{10}$ de probabilité d'erreur à une position nucléotidique donnée de la séquence (tableau 5). Elle sert d'outil pour évaluer la spécificité et la sensibilité théoriques de la méthode ainsi que l'exactitude théorique de l'appel de chacune des bases sur l'ensemble de la séquence interrogée lors d'un test de séquençage. Elle tient compte notamment de l'intensité du signal, du bruit de fond et du rapport entre le signal et le bruit de fond [Ellard *et al.*, 2016; CLSI, 2014; ANPGM, 2012].

Tableau 5 Correspondance entre la valeur de qualité et la probabilité d'erreur de l'appel de la base à une position donnée de la séquence

VALEUR DE QUALITÉ	PROBABILITÉ D'UN APPEL DE BASE ERRONÉ	CERTITUDE DE L'APPEL DE LA BASE
10	1 / 10 (10 %)	90 %
20	1 / 100 (1 %)	99 %
30	1 / 1 000 (0,1 %)	99,9 %
40	1 / 10 000 (0,01 %)	99,99 %
50	1 / 100 000 (0,001 %)	99,999 %

Source: Informations tirées et adaptées de celles publiées par le Clinical and Laboratory Standard Institute [CLSI, 2014].

7.2.2. Couverture de la séquence ciblée

La couverture verticale et la couverture horizontale sont des critères de qualité essentiels à l'appréciation de la qualité d'un essai de SNG et de l'appel d'une variation nucléotidique.

La couverture verticale, aussi appelée profondeur de lecture, correspond au nombre de lectures uniques chevauchantes par position nucléotidique sur la séquence de référence (c.-à-d. le nombre de fois qu'un nucléotide a été séquencé). Elle est exprimée en nombre de « X ». La couverture horizontale correspond à la fraction de la séquence de référence couverte par au moins une lecture. Elle est souvent exprimée en pourcentage (%) des régions d'intérêt séquencées et pour lequel la profondeur de lecture atteint ou dépasse une valeur minimale prédéfinie [Jennings *et al.*, 2017; Gargis *et al.*, 2012].

La performance des outils bio-informatiques des plateformes de SNG pour la détection sensible et spécifique des variations de séquence repose, entre autres, sur ces deux dimensions de la couverture. En effet, la profondeur de lecture d'un nucléotide permet d'évaluer la reproductibilité de la détection d'une base nucléotidique à partir d'un échantillon donné. Un nombre plus élevé de lectures de haute qualité permet donc d'avoir confiance que la base appelée est la bonne, et ce, que celle-ci soit identique (aucune variation nucléotidique détectée) à la base de référence ou différente (variation nucléotidique détectée) de celle-ci [Jennings *et al.*, 2017].

Cependant, plusieurs facteurs influencent la couverture minimale requise pour une région ou une base nucléotidique donnée, dont le type de plateforme, la complexité de la séquence (régions homologues, séquences répétitives, pseudogènes ou régions riches en GC), la qualité de l'ADN utilisé pour l'enrichissement des régions ciblées et les types de variations nucléotidiques recherchées. Ainsi, les valeurs de qualité minimale de couverture verticale et horizontale de chaque test de SNG doivent être systématiquement évaluées lors du développement et de la validation analytique d'un test [Jennings *et al.*, 2017; Gargis *et al.*, 2012].

La profondeur de couverture minimale nécessaire peut être estimée en fonction de la limite de détection requise, de la qualité des lectures et de la tolérance aux résultats faussement positifs ou faussement négatifs. En général, une profondeur de lecture plus faible est acceptable pour les tests constitutionnels, dans lesquels les altérations de la lignée germinale sont plus facilement détectables, car elles se trouvent dans un état hétérozygote ou homozygote. Cependant, dans le cadre de l'analyse de tissus somatiques ou d'ADN mitochondrial, la présence de mosaïcisme peut compliquer l'interprétation de la présence (ou de l'absence) d'une variation nucléotidique [Jennings *et al.*, 2017].

De ce fait, lors de l'analyse de tissu tumoral par séquençage, le laboratoire doit préalablement déterminer la limite de détection (LOD, de l'anglais *limit of detection*) [MM01-A3, 2012] nécessaire afin de garantir une sensibilité adéquate et de minimiser les risques de résultats faussement négatifs. Le pourcentage de cellules tumorales présentes dans l'échantillon analysé doit également être documenté⁶ (voir la [section 7.4](#)).

En ce sens, dans le cas d'une recherche de variations somatiques, l'International Quality Network précise que la profondeur de lecture minimale requise dépend de la proportion de cellules cancéreuses de l'échantillon et de la fréquence allélique de la variation nucléotidique recherchée [Deans *et al.*, 2017]. À titre d'exemple, un échantillon contenant 40 % de cellules cancéreuses, analysé à l'aide d'un test ayant un seuil de détection de 10 % de fréquence allélique, demandera une profondeur de lecture bien plus élevée qu'un échantillon contenant 90 % de cellules cancéreuses pour lequel on vise la détection de variations nucléotidiques jusqu'à un seuil de 30 % de fréquence allélique [Deans *et al.*, 2017].

⁶ Information fournie par un des experts consultés.

À titre informatif, le tableau 6 présente les valeurs de qualité minimales de l'appel des bases et de profondeur de lecture recommandées par certains organismes professionnels et réglementaires pour juger de la qualité d'un essai de SNG et de l'appel d'une variation nucléotidique.

Tableau 6 Recommandations d'organismes professionnels et réglementaires concernant les valeurs de qualité minimales de l'appel des bases et de profondeur de lecture pour l'appel d'une variation nucléotidique

ORGANISME	QUALITÉ DE L'APPEL DES BASES	PROFONDEUR DE LECTURE
Recherche de variations pathogéniques héréditaires		
AMP & CAP [Jennings <i>et al.</i> , 2017]	n. d.	≥ 30X, avec une représentativité égale des deux brins d'ADN (sens et antisens)
EuroGentest [Matthijs <i>et al.</i> , 2016]	QV ≥ 20	n. d.
FDA ^a [FDA, 2016]	QV ≥ 30	Pour un variant hétérozygote de lignée germinale d'un panel ciblé : ▪ ≥ 20X de profondeur minimale ▪ ≥ 300X^b de profondeur moyenne (pour 100 % des bases analysées par un panel ciblé ou 97 % pour le génome complet)
État de New York [NY State, 2015]	QV ≥ 20	≥ 10X pour l'appel d'une variation nucléotidique Toute variation de séquence doit être confirmée par une méthode de référence indépendante pour être rapportée. La confirmation des VSN n'est pas nécessaire si la profondeur de lecture est ≥ 100X.
CLSI [CLSI, 2012a]	n. d.	Moyenne : plusieurs centaines Minimale : ≥ 30X
Recherche de variations pathogéniques somatiques		
État de New York [NY State, 2018]	QV ≥ 20	≥ 500X pour l'appel d'une variation nucléotidique
AMP & CAP [Jennings <i>et al.</i> , 2017]	n. d.	▪ ≥ 250X de profondeur minimale par cible^c ▪ ≥ 1 000X de profondeur moyenne pour les variations hétérogènes dans des tissus de faible cellularité tumorale ▪ ≥ 5 000X pour la détection de variations hétéroplasmiques dans l'ADN mitochondrial
IQN Path [Deans <i>et al.</i> , 2017]	n. d.	Tumeurs solides : 300X à 500X par cible

ORGANISME	QUALITÉ DE L'APPEL DES BASES	PROFONDEUR DE LECTURE
INCa [INCa, 2016]	n. d.	Pour une limite de détection de 10 % d'allèles mutés : ▪ $\geq 300X$ pour les protocoles en PCR ▪ $\geq 200X$ pour les protocoles en capture

Abréviations : AMP : Association for Molecular Pathology; CAP : College of American Pathologists; FDA : Food and Drug Administration (États-Unis); INCa : Institut national du cancer (France); IQN Path : International Quality Network for Pathology; n. d. : donnée non disponible; QV : valeur de qualité; VSN : variation d'un seul nucléotide.

^a Version préliminaire (*draft*) distribuée uniquement à des fins de commentaires et de discussion [FDA, 2016]. La version finale et officielle de ce guide, publié le 13 avril 2018, ne présente aucune limite définie concernant la profondeur de couverture, mais précise que les seuils moyens et minimaux de profondeur de couverture doivent être prédéfinis et rapportés sur la base des indications d'utilisation du test et justifiés à l'aide de preuves objectives et de techniques statistiques valides, qui doivent également être rapportées [FDA, 2018].

^b Une profondeur de séquençage $\geq 200X$ n'est nécessaire que pour les variations présumées présentes à une faible fréquence allélique (ex. : variation mitochondriale, postzygotique, tumorale, etc.). Information fournie par un des experts consultés.

^c Dans certaines circonstances, une profondeur minimale de lecture de moins de 250X peut être acceptable, toutefois la pertinence devrait être justifiée en fonction de la limite de détection prévue, de la qualité des lectures et de la tolérance aux résultats faussement positifs ou faussement négatifs [Jennings *et al.*, 2017].

7.3. Taux d'erreur des technologies de SNG

Jusqu'à tout récemment, une des principales limitations du SNG était la fréquence élevée de bases incorrectement identifiées en raison des artéfacts introduits lors de la conservation, de la préparation ou du séquençage des échantillons. Par exemple, les biais d'amplification pendant la PCR peuvent fausser l'interprétation des populations alléliques. De plus, les erreurs introduites par la polymérase, telles que les mauvaises incorporations de base et les réarrangements attribuables à un changement du gabarit d'acides nucléiques (de l'anglais *template switching*), peuvent entraîner des appels de variations nucléotidiques incorrects. En outre, les erreurs peuvent apparaître lors de l'amplification de la région d'intérêt, des cycles de séquençage et de l'analyse de l'image, ce qui signifie qu'approximativement 0,1 % à 1 % des bases seraient identifiées incorrectement [Fox *et al.*, 2014]. Un schéma des principales étapes des techniques typiques de SNG, y compris leurs sources d'erreurs potentielles, est proposé à la [figure B-1](#) de l'annexe B [Ma *et al.*, 2019].

Toutefois, afin de réduire la fréquence globale des erreurs produites par SNG, certaines stratégies utilisant l'étiquetage d'une molécule unique d'ADN à l'aide de codes à barres (de l'anglais *tag-based methods*) ont été proposées. En outre, le séquençage duplex (de l'anglais *duplex sequencing*) utilise l'étiquetage aléatoire de l'ADN double brin à l'aide de séquences dégénérées en plus d'adaptateurs permettant de reconnaître les lectures provenant de chacun des deux brins. Il est donc possible de réduire significativement les erreurs de séquençage en comparant les séquences dérivées d'un brin d'ADN avec celles de son complément. Il a été rapporté que cette méthode pourrait réduire les erreurs à moins d'une pour dix millions de nucléotides séquencés [Fox *et al.*, 2014].

Par ailleurs, l'utilisation de filtres *in silico* de suppression des erreurs permettrait de réduire le taux de substitutions, type d'erreur fréquemment rencontré en séquençage, à un niveau 10 à 100 fois inférieur à celui généralement accepté pour les techniques de SNG [Ma *et al.*, 2019].

Le tableau 7 présente la fréquence et les types d'erreurs de différentes technologies de séquençage. Précisons que la quantité et le type d'erreurs produites en cours de séquençage sont grandement dépendants de l'instrument et de la chimie utilisés. De ce fait, les taux d'erreur des différentes plateformes de SNG ne sont pas nécessairement comparables [Gargis *et al.*, 2012; Glenn, 2011].

Tableau 7 Fréquence et types d'erreurs de différentes technologies et plateformes de séquençage

PLATEFORMES COMMERCIALES	MÉTHODE	TYPE D'ERREUR LE PLUS FRÉQUENT	FRÉQUENCE D'ERREURS PAR BASE
Sanger 3730xl (capillaire)	Terminaison de chaîne	Substitutions d'un nucléotide	10^{-2}
454 GS Junior	Pyroséquençage	Délétions	10^{-2}
PacBio RS	Molécule unique en temps réel (sans amplification)	Délétions CG	10^{-2}
Ion Torrent PGM	Détection d'ions hydrogène (pH)	Courtes délétions	10^{-2}
Solid	Ligation et incorporation à deux bases	Biais A-T	2×10^{-2}
Illumina MiSeq	Terminateurs réversibles	Substitutions d'un nucléotide	10^{-3}
Illumina HiSeq			
Méthodes de correction d'erreur <i>Tag-based</i>			
SafeSeq	Code à barres moléculaire	Substitutions d'un nucléotide	$1,4 \times 10^{-5}$
CircleSeq	Multiples copies de molécules d'ADN circularisés	Substitutions d'un nucléotide	$7,6 \times 10^{-6}$
Duplex Sequencing	Comparaison des séquences des brins complémentaires	Substitutions d'un nucléotide	5×10^{-8}

Source : Fox *et al.*, 2014.

Considérant les nouvelles performances de certaines techniques de SNG, la norme actuelle voulant que les variations nucléotidiques détectées par SNG soient validées par séquençage Sanger a été remise en question par le National Institute of Health (NIH) du NISC Comparative Sequencing Program. Les travaux de ce groupe ont montré que le reséquençage Sanger de 5 660 variations nucléotidiques détectées par SNG a présenté un taux de concordance de 99,965 % (IC95 % de 99,887 à 99,993 %) [Beck *et al.*, 2016].

De plus, selon l'analyse des résultats discordants, un seul cycle de séquençage Sanger serait plus susceptible de réfuter incorrectement un vrai positif par SNG que d'identifier correctement un faux positif par SNG. Pour ces raisons, les auteurs considèrent que la validation par séquençage Sanger des variations nucléotidiques obtenues par SNG présente une utilité limitée [Beck *et al.*, 2016]. Cependant, les organisations internationales et celles produisant les règles de bonne pratique n'ont pas encore toutes adopté cette vision.

7.4. Critères de performance à évaluer lors de la validation analytique d'un test effectué par SNG

La validation analytique d'un test clinique comporte différentes dimensions qui varient en fonction du type d'analyse (quantitative, qualitative ou semi-quantitative), de la technologie utilisée, de la nature de la condition ou de la maladie (héréditaire ou acquise, prévalence) et de la stringence du test (test de dépistage, de diagnostic ou de confirmation) [CLSI, 2012a; Mattocks *et al.*, 2010]. Le tableau 8 présente les différentes dimensions de performance devant être mesurées pour la validation analytique d'un test clinique, ainsi que leur définition adaptée au séquençage. Une description plus exhaustive des termes est également fournie à l'[annexe D](#).

Tableau 8 Définitions des critères de performance à évaluer lors de la procédure de validation analytique d'un test clinique effectué par SNG

CRITÈRE DE PERFORMANCE	DÉFINITION*	PARAMÈTRES ET CONSIDÉRATIONS
Exactitude	Degré de concordance entre la séquence d'acide nucléotidique déterminée par le test et la séquence de référence. $VP + VN / (VP + FP + VN + FN)$	Valeur de qualité de l'appel des bases, couverture, couverture moyenne, uniformité de la couverture, fraction allélique, duplication des lectures, biais associé au GC et au brin d'ADN [Gargis <i>et al.</i> , 2012].
Sensibilité	Probabilité qu'un test détecte une variation nucléotidique lorsqu'elle est présente. $VP / (VP + FN)$	Peut varier en fonction de la couverture, du type de variation nucléotidique, du contexte et de la qualité de l'appel des bases. La sensibilité doit être évaluée d'après un seuil de couverture donné pour les régions du génome ciblées par l'analyse [Gargis <i>et al.</i> , 2012].
Spécificité	Probabilité qu'un test ne détecte pas de variation nucléotidique lorsqu'aucune n'est présente. $VN / (VN + FP)$	Peut varier en fonction de la couverture, du type de variant, du contexte de la séquence et de la qualité de l'appel des bases. La spécificité doit être évaluée d'après un seuil de couverture donné pour les régions du génome ciblées par l'analyse [Gargis <i>et al.</i> , 2012].
Précision	Degré avec lequel des analyses répétées d'un même échantillon donnent le même résultat. Répétabilité (précision intra-essai) : Degré avec lequel la même séquence est obtenue lorsqu'un même échantillon est séquençé à de multiples reprises sous les mêmes conditions.	Valeur de qualité de l'appel des bases, couverture, uniformité de la couverture et ratio transition/transversion doivent être constants entre les essais et respecter les critères de qualité minimaux établis lors de la validation de la méthode [Gargis <i>et al.</i> , 2012].

CRITÈRE DE PERFORMANCE	DÉFINITION*	PARAMÈTRES ET CONSIDÉRATIONS
	Fidélité intermédiaire (précision inter-essai) : degré avec lequel la même séquence est obtenue lorsqu'un même échantillon est séquencé à de multiples reprises par différents opérateurs, utilisant différents lots de réactifs et sur une période plus longue que pour l'estimation de la répétabilité. Reproductibilité (précision inter-laboratoire) : degré avec lequel la même séquence est obtenue lorsqu'un même échantillon est séquencé par différents laboratoires.	
Limite de détection	La fraction allélique minimale détectable dans un échantillon donné. La plus faible valeur limite (pourcentage de mutants) permettant de distinguer un signal muté d'un signal non muté[†].	Applicable dans les cas de variations de séquences minoritaires tels l'hétéroplasmie mitochondriale, le mosaïcisme, les échantillons tumoraux et la détection prénatale des aneuploïdies.
Intervalle de mesure	Région du génome pour laquelle une séquence de qualité acceptable peut être obtenue par le test.	Peut inclure la région interne d'un gène interrogé, la région codante ou d'autres régions comportant un nombre limité de gènes, le statut homozygote ou hétérozygote d'une mutation.
Intervalle de référence	Gamme des variations nucléotidiques détectées dans une population non affectée. Les résultats en dehors de cet intervalle pourraient être cliniquement significatifs.	Séquence bien caractérisée de la région du gène ou du génome interrogé, incluant notamment la distribution des variations de séquences et, le cas échéant, des variations pathogéniques connues.

Abréviations : FN : faux négatifs; FP : faux positifs; VN : vrais négatifs; VP : vrais positifs.

* Définitions tirées et/ou inspirées de celles publiées par le groupe de travail Nex-StoCT, formé par le Center for Disease Control and Prevention (CDC) aux États-Unis [Gargis *et al.*, 2012], sauf si autrement spécifié. Ces définitions sont également reprises par le CLSI [CLSI, 2014].

† Définition publiée par l'Institut national du cancer (INCa) en France [INCa, 2014].

7.5. Détermination du nombre de variations nucléotidiques et d'échantillons cliniques à analyser pour l'évaluation de la validité analytique d'un test effectué par SNG

Divers organismes, tels que le Conseil canadien des normes, le College of American Pathologists (CAP), le Clinical and Laboratory Standards Institute (CLSI), l'American College of Medical Genetics and Genomics (ACMG) ainsi que certaines agences réglementaires américaines et européennes, requièrent ou recommandent que les laboratoires cliniques établissent ou vérifient la sensibilité, la spécificité, l'exactitude, la précision ainsi que plusieurs autres caractéristiques de performance du test. L'[annexe E](#) présente divers organismes s'étant prononcés sur les caractéristiques de performance à valider dans le contexte des tests de SNG ou, de façon plus générale, dans le contexte des analyses génétiques et moléculaires à visée clinique.

Précisons que le nombre total d'échantillons cliniques requis pour la validation d'un test par SNG dépend de nombreux facteurs, notamment :

- l'application clinique du test (dépistage, diagnostic, confirmation);
- l'ampleur de la ou des régions du génome analysées;
- le nombre d'analytes spécifiques que le test est censé détecter (différents types de variations nucléotidiques);
- le nombre de variations nucléotidiques de différents types que comportent les échantillons;
- les types d'échantillons acceptés pour le test;
- la disponibilité des échantillons cliniques, particulièrement des échantillons anormaux comportant des variations pathogéniques connues;
- l'expérience acquise par le laboratoire relativement à la technique;
- le coût et la durée des analyses.

De ce fait, bien qu'ils insistent sur la nécessité de valider un test clinique conçu en laboratoire, plusieurs organismes ne se sont pas prononcés sur un nombre minimal défini d'échantillons requis pour la validation d'un test clinique par SNG. À cet égard, le College of American Pathologists (CAP) conclut que les considérations statistiques concernant le nombre d'échantillons nécessaires à la validation d'un test ne peuvent pas être universellement appliquées aux divers et nombreux essais potentiellement effectués par SNG [Aziz *et al.*, 2015].

Le tableau 9 présente néanmoins quelques organismes qui ont publié des recommandations quant au nombre d'échantillons requis pour la validation de la méthode ou d'un test génétique par SNG (ou par d'autres techniques moléculaires). Selon la littérature, le nombre d'échantillons utilisés pour la validation de diverses applications du SNG varie de 20 à plus de 80 échantillons [Aziz *et al.*, 2015]. Dans l'éventualité où une variation pathogénique connue d'un échantillon témoin ne serait pas détectée, le test doit être modifié et le processus de validation, répété [Deans *et al.*, 2017].

Précisons que, dans le cadre des analyses effectuées par SNG, où plusieurs variations nucléotidiques peuvent être détectées chez un même patient, la notion de vrais ou faux positifs est plutôt définie en fonction du nombre de variations nucléotidiques détectées et non en fonction du nombre d'échantillons analysés [INCa, 2016]. Le nombre d'échantillons nécessaires pour effectuer une validation de puissance statistique suffisante peut être diminué en fonction du nombre de mutations comprises dans ces échantillons [INCa, 2016]. De plus, comme il a été mentionné précédemment, bien que des références commerciales contenant des centaines de variations de séquence connues et de tous les types peuvent être utilisées dans les études de validation d'un nouveau test, une étude de concordance à l'aide d'un certain nombre d'échantillons cliniques, analysés en parallèle par un autre laboratoire ou par une autre technique, doit également être effectuée (tableau 9).

Dans le cas particulier des maladies rares, pour lesquelles l'accessibilité des échantillons cliniques pertinents est difficile, l'accumulation de données dans le temps est un aspect important de la validation initiale d'un test, laquelle peut être utilisée pour l'évaluation continue de l'exactitude et de la qualité du test [Mattocks *et al.*, 2010; Jennings *et al.*, 2009; Maddalena *et al.*, 2005].

Tableau 9 Recommandations d'organismes professionnels et réglementaires concernant le nombre minimal de variations nucléotidiques ou d'échantillons cliniques requis pour la validation analytique d'un test clinique par SNG

ORGANISME [AUTEUR, ANNÉE]	EXACTITUDE, SENSIBILITÉ ET SPÉCIFICITÉ	PRÉCISION		ÉTUDE DE CONCORDANCE
		INTRA-ESSAI (RÉPÉTABILITÉ)	INTER-ESSAI (REPRODUCTIBILITÉ)	
Recherche de variations nucléotidiques héréditaires				
CLSI	30 échantillons (MM01-A3) 100 échantillons sur une période de 10 à 20 jours (EP12-A2), soit 50 sauvages et 50 avec variations nucléotidiques	n. d.	Plusieurs en 3 essais (jours, réactifs et opérateurs différents) (MM01-A3)	30 échantillons (MM01-A3)
CCMG [Hume <i>et al.</i> , 2019]	≥ 60 variations nucléotidiques au total, dont au moins 10 par type de variations ^a	n. d.	n. d.	n. d.
EuroGentest [Matthijs <i>et al.</i> , 2016]	~ 300 variations d'un échantillon de référence commercial (ex. : NIST) ou par validation orthogonale (Sanger ou autre méthode SNG)	≥ 3 échantillons en triplicata	3 échantillons analysés par des essais indépendants sous des différentes conditions	n. d.
ACGS [Deans <i>et al.</i> , 2015]	Essentiel : ≥ 60 variations uniques Recommandé : ≥ 150 variations uniques	n. d.	Plusieurs échantillons analysés par ≥ 3 essais indépendants	n. d.
État de New York [NY State, 2015]	1 échantillon de référence commercial (ex. : Hapmap ou NIST)	n. d.	≥ 3 échantillons analysés simultanément par 3 essais indépendants	25 patients divisés pour être analysés par 3 essais indépendants
ACMG [Rehm <i>et al.</i> , 2013]	≥ 1 échantillon de référence commercial (ex. : Corriel ou NIST)	1 échantillon en duplicata ou en triplicata	1 échantillon analysé par 2 ou 3 essais indépendants	Variété d'échantillons porteurs de différents types de variations génétiques, analysés à l'aveugle
Recherche de variations nucléotidiques somatiques				
État de New York ^b [NY State, 2018]	▪ ≥ 3 échantillons de référence commerciaux (ex. : Hapmap NA12878, NA19240 ou NIST) et le taux d'erreur devrait être < 2 % ▪ ≥ 50 échantillons cliniques comprenant tous les types de tumeurs et de spécimens cliniques qui seront acceptés pour l'analyse	≥ 3 échantillons positifs en triplicata de chaque type de variations nucléotidiques	4 à 5 échantillons positifs pour chaque type de variations nucléotidiques en 4 à 5 essais	10 échantillons positifs de tous les types de variations que le test propose de détecter dans chaque région ciblée par le test

ORGANISME [AUTEUR, ANNÉE]	EXACTITUDE, SENSIBILITÉ ET SPÉCIFICITÉ	PRÉCISION		ÉTUDE DE CONCORDANCE
		INTRA-ESSAI (RÉPÉTABILITÉ)	INTER-ESSAI (REPRODUCTIBILITÉ)	
AMP & CAP [Jennings <i>et al.</i> , 2017]	▪ ≥ 2 échantillons de référence (ex. : HapMap) ▪ ≥ 59 échantillons pour évaluer les paramètres de qualité et de performance du test^c	n. d.	≥ 3 échantillons à toutes les étapes du SNG, afin d'inclure différents instruments, opérateurs et lots de réactifs	n. d.
INCa [INCa, 2016]	30 échantillons (15 mutés et 15 normaux) ^d	1 échantillon en triplicata	30 échantillons sur 15 jours (2 niveaux d'allèles mutés)	n. d.
PSGBI [Salto-Tellez et Gonzalez de Castro, 2014]	≥ 300 variations (pour un panel de 20 kb)	n. d.	n. d.	n. d.

Abréviations : ACGS : Association for Clinical Genetic Science (Royaume-Uni); ACMG : American College of Medical Genetics and Genomics (États-Unis); APP : accord de pourcentage positif [vrais positifs/(vrais positifs + faux négatifs)]; ASHI : American Society for Histocompatibility and Immunogenetics (États-Unis); CAP : College of American Pathologists (États-Unis); CCMG : Canadian College of Medical Geneticist; CLSI : Clinical & Laboratory Standards Institute (États-Unis); INCa : Institut national du cancer (France); kb : kilobase; PSGBI : Pathological Society of Great Britain and Ireland; VPP : valeur prédictive positive [vrais positifs/(vrais positifs + faux positifs)].

^a Si la cible ne peut être atteinte lors de la validation (ex. : rareté de certains types de variations pathogéniques), la vérification de variations nucléotidiques additionnelles devrait se poursuivre lors de l'implantation jusqu'à ce que la cible soit atteinte [Hume *et al.*, 2019].

^b Le Department of Health de l'État de New York propose un protocole complet pour le développement et la validation analytique de tests cliniques effectués par SNG pour la détection de variations somatiques. Des nombres précis d'échantillons cliniques exigés pour la validation de différents types de variations nucléotidiques y sont décrits [NY State, 2018].

^c Incluant, dans la mesure du possible, des échantillons cliniques préalablement caractérisés, du même type que ceux destinés au dosage (ex. : sang, moelle, tissus FFEP), représentant chaque type de variations pathogéniques que le test prétend pouvoir détecter (ex. : VSN, VMN, translocation, etc.) ainsi que les variations nucléotidiques les plus courantes et les plus pertinentes selon l'utilisation prévue du test [Jennings *et al.*, 2017].

^d La notion de vrais et de faux positifs se définissant en fonction du nombre de variations de séquence, le nombre d'échantillons cliniques peut être diminué en fonction du nombre de variations comprises dans ces échantillons [INCa, 2016].

7.6. Estimation de la puissance statistique d'une étude de validation analytique en fonction de la taille de l'échantillon

La détermination de la puissance statistique d'une étude de validation par des calculs précis peut être complexe. Pour des raisons pratiques, Mattocks et ses collaborateurs [2010] proposent une méthode simple afin d'estimer la puissance statistique d'une étude de validation en fonction de la taille de l'échantillon. Selon les auteurs, cette façon de faire procure une estimation suffisamment précise de la puissance statistique pour les tests de type qualitatif.

L'objectif de l'exercice est de déterminer le nombre de mesures qui procurera une puissance statistique suffisante pour déterminer la sensibilité et la spécificité d'un test, au degré de confiance désiré, pour une application donnée [Mattocks *et al.*, 2010].

Selon cette règle, pour un intervalle de confiance (IC) à 95 %, la probabilité qu'un événement passe inaperçu lors de l'étude de validation d'un nombre « N » de mesures est de « 3/N ». Les auteurs illustrent ce principe à l'aide des exemples du tableau 10. Comme la sensibilité des technologies utilisées en génétique moléculaire est généralement élevée, les auteurs émettent l'hypothèse que le test offre une sensibilité observée de 100 %.

Tableau 10 Effet de la taille de l'échantillon sur la puissance statistique lors de la détermination de la sensibilité d'un test qualitatif

N	SENSIBILITÉ OBSERVÉE (%)	3/N (PROBABILITÉ D'UN FAUX NÉGATIF)	SENSIBILITÉ RÉELLE PROBABLE (%), IC À 95 %
30	30/30 (100)	0,1 (10 %)	≥ 90
300	300/300 (100)	0,01 (1 %)	≥ 99

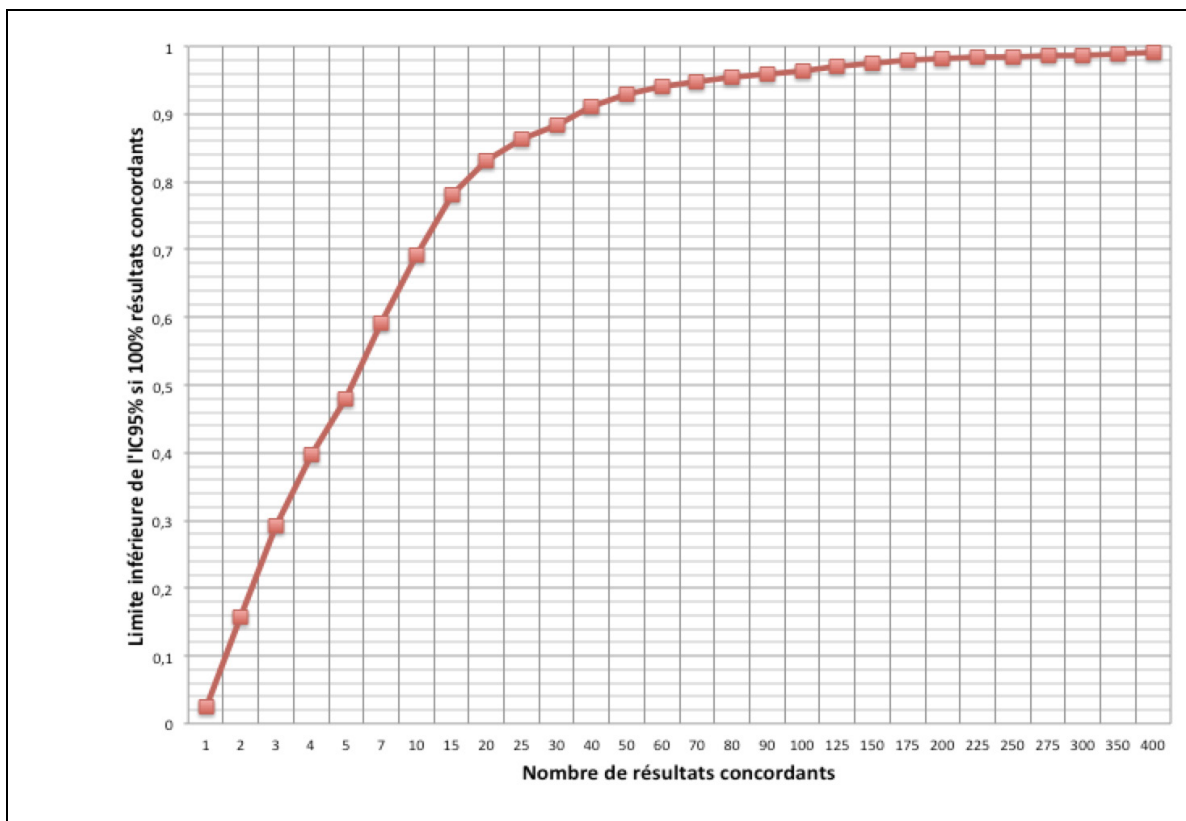
Source : Mattocks et ses collaborateurs [2010].

Abréviation : N : nombre d'échantillons.

Précisons que cette règle peut s'appliquer à tous les calculs de proportion et peut donc être utilisée pour déterminer la spécificité autant que la sensibilité. La proportion de résultats positifs (présence d'une variation nucléotidique) par rapport aux sujets connus positifs déterminera la fiabilité de l'estimation de la sensibilité, alors que la proportion de résultats négatifs (résultats normaux ou absence de variation nucléotidique) par rapport aux sujets normaux déterminera celle de la spécificité [Mattocks *et al.*, 2010].

Les auteurs soulignent que cette méthode procurera une estimation précise pour les études comportant un minimum de 60 mesures pour chaque proportion estimée [Mattocks *et al.*, 2010]. La figure 2 présente l'incertitude associée à la puissance de l'estimation d'une proportion (limite inférieure de l'intervalle de confiance à 95 %) lorsque 100 % des résultats sont concordants.

Figure 2 Incertitude de l'estimation d'une proportion de 100 % en fonction du nombre d'observations



Abréviation : IC95 % : intervalle de confiance à 95 %.

Cet outil procure des estimations utiles du nombre de mesures requises pour une validation analytique de puissance statistique satisfaisante.

L'analyse d'un nombre suffisant de variations nucléotidiques pour atteindre une puissance statistique satisfaisante est relativement facile avec des échantillons de référence commerciaux (ex. : Corriel, HapMap, NIST). Toutefois, la disponibilité des échantillons cliniques appropriés, nécessaires pour évaluer les paramètres de qualité et de performance d'un test, notamment lors des études de concordance, est souvent un facteur limitant [ANPGM, 2012; Mattocks *et al.*, 2010; Jennings *et al.*, 2009; Maddalena *et al.*, 2005].

En effet, selon Jennings et ses collaborateurs [2009], des données statistiquement significatives ne peuvent pas nécessairement être obtenues pour tous les aspects de chaque étude de validation. Selon les auteurs, il n'y a pas toujours suffisamment d'échantillons cliniques, de temps et de soutien financier, scientifique ou technologique pour permettre une telle collecte de données pour chaque test diagnostique [Jennings *et al.*, 2009]. Le CLSI (MM20) n'adhère cependant pas à cette vision de baisser les standards de validation rigoureuse des tests diagnostiques moléculaires, car ils ne seront souvent réalisés qu'une seule fois dans la vie d'un individu (pour les maladies héréditaires) et conditionneront ses soins de santé pour le reste de sa vie, sans parler des risques de discrimination et de stigmatisation.

Par contre, pour augmenter la puissance statistique d'une étude de validation, les spécimens cliniques obtenus localement peuvent être complétés avec ceux provenant d'autres laboratoires et ceux disponibles commercialement [Peris-Vicente *et al.*, 2014; Pont-Kingdon *et al.*, 2012; Mattocks *et al.*, 2010]. Le matériel résiduel provenant de spécimens reçus pour l'analyse clinique d'autres conditions ou maladies peut parfois servir d'échantillons témoins négatifs. Ces échantillons doivent être anonymisés avant d'être utilisés [Ellard *et al.*, 2016; Pont-Kingdon *et al.*, 2012].

De plus, dans le cas particulier des maladies rares, pour lesquelles l'accessibilité des échantillons cliniques positifs est difficile, l'accumulation de données dans le temps est un aspect important de la validation initiale d'un test, laquelle peut être utilisée pour l'amélioration continue de l'évaluation de l'exactitude et de la qualité du test [Mattocks *et al.*, 2010; Jennings *et al.*, 2009; Maddalena *et al.*, 2005].

8. ÉTAPES POSTANALYTIQUES

8.1. Interprétation des variations nucléotidiques héréditaires et communication des données de séquence

L'American College of Medical Genetics and Genomics (ACMG), l'Association for Molecular Pathology (AMP) et le College of American Pathologists (CAP) ont conjointement publié des standards et des recommandations pour l'interprétation des variations nucléotidiques [Richards *et al.*, 2015]. Les principaux aspects de ces recommandations sont présentés ici. Ils sont complétés par certaines recommandations de l'Association for Clinical Genetic Science (ACGS) [Deans *et al.*, 2015] et du Canadian College of Medical Geneticists (CCMG) [Hume *et al.*, 2019].

- Le processus de classification des variations pathogéniques des désordres héréditaires est effectué en fonction de divers éléments de preuve (ex. : données populationnelles et démographiques, données informatiques *in silico*, données fonctionnelles, données de ségrégation, etc.). Ce processus permet de classer les variations nucléotidiques en cinq catégories correspondant aux terminologies suivantes [Richards *et al.*, 2015] :
 - Variation pathogénique;
 - Variation nucléotidique probablement pathogénique (certitude de plus de 90 % que le variant est pathogène);
 - Variation nucléotidique de signification incertaine;
 - Variation nucléotidique probablement bénigne (certitude de plus de 90 % que la variation nucléotidique est bénigne);
 - Variation nucléotidique bénigne.
- Le rapport devrait contenir l'information exacte concernant la nature et le positionnement de la variation nucléotidique en comparaison de la séquence de référence, de même que l'impact de ce changement sur la séquence de la protéine, et ce, en accord avec la nomenclature standardisée du Human Genome Variation Society (HGVS)⁷. Le logiciel Mutalyzer peut être utilisé dans la plupart des cas pour vérifier la nomenclature des variations nucléotidiques détectées⁸ [Deans *et al.*, 2015; Richards *et al.*, 2015].
- Le rapport devrait également inclure une description de la condition clinique testée (maladie), la liste des gènes ainsi que leur séquence de référence (incluant le numéro d'identification), les exons, l'étendue des régions introniques flanquantes analysées, les informations relatives à l'hérédité, la zygosité, et ce, toujours en

⁷ HGVS nomenclature. Sequence Variant Nomenclature [site Web], disponible à : <http://varnomen.hgvs.org/>.

⁸ Mutalyzer [site Web], disponible à : <https://mutalyzer.nl/>.

accord avec la nomenclature standardisée. L'origine parentale peut également être incluse si elle est connue [Ellard *et al.*, 2016; Richards *et al.*, 2015].

- La méthode d'enrichissement (ex. : élaboration personnalisée des amorces, trousse ou panel commercial) et l'instrument d'analyse devraient être mentionnés [Ellard *et al.*, 2016].
- L'interprétation des résultats devrait contenir les éléments de preuve appuyant la classification des variations nucléotidiques, y compris leur effet prédit sur la protéine résultante, et préciser si les variations nucléotidiques détectées sont susceptibles d'expliquer, en tout ou en partie, la condition clinique du patient et pour laquelle le test est indiqué [Richards *et al.*, 2015].
- Le rapport devrait également inclure des recommandations destinées au clinicien pour des tests cliniques supplémentaires, tels que des tests enzymatiques et fonctionnels à partir des cellules du patient, ainsi que l'intérêt de tester les membres de la famille afin de permettre une interprétation plus approfondie des variations nucléotidiques [Richards *et al.*, 2015].
- Il convient de noter si la variation nucléotidique a été rapportée précédemment dans la littérature ou dans des bases de données [Richards *et al.*, 2015].
- Les résultats négatifs ne doivent être déclarés que pour les nucléotides analysés ayant produit des données de séquence de haute qualité [Ellard *et al.*, 2016].
- L'étendue de l'analyse, y compris les limites du test et les risques potentiels de faux positifs, doit également être déclarée. Les éléments suivants devraient être spécifiés [Deans *et al.*, 2015] :
 - les types de variations nucléotidiques détectables par le test (ex. : SNV, CNV, indels);
 - la taille maximale des insertions et des délétions détectables, le cas échéant;
 - la limite de détection de fréquence allélique (LOD);
 - les limites de la méthode, incluant la détection des variations nucléotidiques localisées dans des régions homologues, des régions répétitives et de celles riches en GC;
 - les méthodes additionnelles utilisées pour compléter l'analyse des régions dont la performance n'était pas suffisante par SNG;
 - la couverture horizontale pour le panel complet et pour les gènes individuels;
 - la profondeur de lecture (couverture verticale) minimale à laquelle les données ont été considérées comme satisfaisantes pour la production du rapport;
 - les régions génomiques ciblées pour lesquelles les critères de performance de l'analyse ont été satisfaits;

- les régions génomiques ciblées pour lesquelles les critères de performance de l'analyse n'ont pas été satisfaits.
- Le CCMG recommande que le laboratoire définisse le ratio allélique parmi les données de séquence auquel une variation nucléotidique est considérée comme hétérozygote ou homozygote [Hume *et al.*, 2019]. Le laboratoire devrait développer une politique interne relative à la détection de variations nucléotidiques à un ratio allélique plus faible qu'attendu, puisqu'elle pourrait correspondre à du mosaïcisme ou à un pseudogène.
- Les régions de séquence ne répondant pas aux critères de qualité requis de couverture (verticale et horizontale) peuvent être décrites comme des régions à faible couverture ou peuvent être omises du rapport clinique. Si une région revêt un intérêt clinique particulier, l'utilisation d'une autre méthode afin de compléter son analyse devrait être envisagée (ex. : séquençage Sanger) [Deans *et al.*, 2015].
- En l'absence d'un système de transfert de tubes robuste et fiable (ex. : balayage de code à barres, transferts avec témoin) assurant l'identité de l'échantillon à chaque étape, les variations détectées pour la première fois dans un pedigree particulier devraient être confirmées par un second test à l'aide d'un échantillon frais [Deans *et al.*, 2015].
- L'ACMG recommande que les résultats soient interprétés par un spécialiste certifié en génétique moléculaire clinique, un médecin de laboratoire formé en génétique moléculaire, ou l'équivalent [Richards *et al.*, 2015].

8.2. Interprétation des variations nucléotidiques somatiques

La description des variations nucléotidiques somatiques, principalement celles observées dans les cellules cancéreuses, comporte une complexité non rencontrée avec les variations constitutionnelles. À titre d'exemple, les rapports entre les allèles sont très variables et l'hétérogénéité tumorale peut provoquer des différences d'échantillonnage. Les variations peuvent présenter une valeur diagnostique, pronostique, thérapeutique ou prédictive de la réponse à un traitement, ce qui complique davantage leur interprétation et leur classification. Cet exercice nécessite notamment des bases de données spécifiques à la tumeur ou au cancer, en plus des bases de données utilisées pour l'interprétation des tissus constitutionnels [Li *et al.*, 2017; Richards *et al.*, 2015].

L'Association for Molecular Pathology (AMP), l'American Society of Clinical Oncology (ASCO) et le College of American Pathologist (CAP) ont récemment publié des recommandations conjointes concernant l'interprétation, l'annotation, la classification et la communication des variations de séquence en oncologie [Li *et al.*, 2017]. Selon ces recommandations, les variations nucléotidiques détectées dans les désordres oncologiques peuvent être classées en quatre catégories en fonction de leur signification clinique dans le diagnostic, le pronostic et la prise en charge thérapeutique de la condition du patient [Li *et al.*, 2017]. Ces catégories sont basées sur le niveau de preuve scientifique de pathogénicité des variations nucléotidiques détectées (tableau 11).

Tableau 11 Catégorisation des variations nucléotidiques somatiques en fonction du niveau de preuve scientifique

NIVEAU	CLASSIFICATION DE LA VARIATION NUCLÉOTIDIQUE	UTILITÉ CLINIQUE	NIVEAU DE PREUVE
I	Signification clinique robuste	Valeur diagnostique, pronostique et thérapeutique	Niveau A : Thérapie approuvée par la FDA. Variant inclus dans les guides cliniques professionnels. Niveau B : Études de qualité avec consensus d'experts dans le domaine.
II	Signification clinique potentielle	Valeur diagnostique, pronostique et thérapeutique	Niveau C : Thérapie approuvée par la FDA pour différents types de tumeurs ou thérapies expérimentales. Multiples études de faible envergure publiées avec un certain consensus. Niveau D : Essais précliniques ou quelques rapports de cas sans consensus.
III	Signification clinique inconnue	Aucune	Variant non observé à une fréquence allélique significative dans les bases de données générales ou spécifiques à la sous-population testée, ou dans les bases de données pancancéreuses ou spécifiques à une tumeur. Aucune preuve convaincante publiée d'association avec le cancer.
IV	Bénin ou probablement bénin	Aucune	Variant observé à une fréquence allélique significative dans les bases de données générales ou spécifiques à la sous-population testée. Aucune preuve publiée d'association avec le cancer.

Source : Li *et al.*, 2017.

Abbréviations : FDA : Food and Drug Administration (États-Unis).

Outre la classification des variations somatiques, les principaux aspects des recommandations proposées sont présentés ici. Ils sont complétés par certaines recommandations du Canadian College of Medical Geneticists (CCMG) [Hume *et al.*, 2019].

- Les variations nucléotidiques détectées doivent être soigneusement examinées par des professionnels en diagnostic moléculaire, et ce, en tenant compte du contexte global de chaque cas clinique, incluant notamment les résultats histologiques. La catégorisation des variations doit être fondée sur des preuves et doit être effectuée avant la création du rapport médical.
- Les altérations génétiques détectées et correspondant aux niveaux I à III doivent être rapportées en ordre décroissant d'importance clinique. Il n'est pas recommandé d'inclure les variations de niveau IV dans le rapport.

- Toute altération génétique détectée doit être annotée et rapportée selon la nomenclature du HUGO Gene Nomenclature Committee (HGNC)⁹ affilié au Human Genome Variation Society (HGVS)¹⁰ :
 - Les SNV et les indels doivent être rapportés en utilisant les annotations p. et c. (ex. : *BRAF* p.V600E, c.1799T>A);
 - Les gènes de fusion doivent être rapportés en identifiant les deux partenaires de fusion séparés par une barre oblique « / » (ex. : fusion *EML4/ALK*);
 - Les CNV générés par une analyse de SNG doivent être rapportés sous forme de tableau comme GAIN ou PERTE (*LOSS*) du nombre de copies. Les coordonnées génomiques du gène ou du locus peuvent être incluses s'il y a lieu. Le rapport des changements du nombre de copies peut être effectué (ex. : GAIN du nombre de copies *EGFR* [nombre de copies 25]; PERTE du nombre de copies *CDKN2A*).
- Au besoin, la nomenclature familière devrait être ajoutée à la nomenclature standard afin de donner un sens clair aux médecins qui lisent les rapports et les utilisent pour déterminer la prise en charge clinique du patient.
- Le rapport devrait également contenir plusieurs autres éléments qui peuvent être pertinents pour une analyse plus approfondie des résultats ou pour la comparaison avec d'autres résultats obtenus du patient au fil du temps, tels que les coordonnées génomiques, la construction du génome et la séquence de référence de transcription utilisée (ex. : NM_004333.4).
- La fraction allélique et la couverture doivent être évaluées et incluses dans le rapport. Le rapport doit également inclure le seuil de couverture du séquençage pour le test de SNG utilisé. Tous les gènes et points chauds ne répondant pas aux critères minimaux de couverture de séquençage prédéfinis doivent être déclarés comme étant invalides.
- Les rapports ne devraient pas nécessairement se limiter aux résultats positifs. Les résultats négatifs pertinents doivent être rapportés, d'une manière spécifique à la maladie ou à la condition du patient.
- Les incertitudes, si elles sont présentes, doivent être communiquées dans le rapport, notamment les problèmes de qualité de la séquence, de l'adéquation de l'échantillon, du contenu tumoral et des connaissances biomédicales.
- Si la détection d'une variation pathogénique à une fréquence allélique suggère un statut hétérozygote ou homozygote dans un gène associé à une prédisposition héréditaire, le rapport devrait recommander d'orienter le patient vers une

⁹ HUGO Gene Nomenclature Committee (HGNC) : <https://www.genenames.org/>.

¹⁰ Human Genome Variation Society (HGVS) : <http://www.hgvs.org/>.

consultation en génétique afin d'évaluer la possibilité d'un syndrome héréditaire [Hume *et al.*, 2019].

8.3. Assurance de la qualité et épreuves de compétence

L'évaluation continue de la qualité des données et des résultats cliniques de SNG constitue une norme établie. Les laboratoires cliniques offrant des services de SNG doivent donc se doter d'un programme d'assurance de la qualité [Rehm *et al.*, 2013; Gargis *et al.*, 2012]. Ce programme doit entre autres comporter des méthodes de contrôle interne et externe de la qualité de leur processus de SNG.

8.3.1. Contrôles internes de la qualité

Afin de s'assurer de la qualité d'une analyse effectuée par SNG, il est généralement nécessaire d'inclure des contrôles internes de qualité, soit un contrôle négatif (absence d'échantillon ou blanc) et un contrôle positif (échantillon de référence connu et validé). Ceux-ci doivent être introduits au tout début du processus de SNG et subir les mêmes étapes de manipulation que l'échantillon clinique, de la quantification initiale au traitement final des données en aval.

L'analyse simultanée de ces contrôles internes permet la surveillance en temps réel de la contamination et de la qualité des données à certaines étapes du flux de travail de SNG. Ces étapes comprennent généralement l'évaluation initiale de l'échantillon, la fragmentation, l'évaluation finale de la librairie, le suivi des taux d'erreur pendant le processus de séquençage et l'analyse des données brutes en mettant l'accent sur la qualité de lecture [Rehm *et al.*, 2013].

Précisons qu'un contrôle positif ne teste pas chaque variante possiblement détectée par SNG. Un laboratoire peut posséder plusieurs contrôles positifs, comportant une gamme de variations de séquence, qui seront utilisés en rotation afin d'inclure les variantes pathogéniques majeures ou fréquentes dans le type d'échantillons cliniques analysés.

8.3.2. Contrôles externes de la qualité

Les laboratoires qui utilisent des méthodes de séquençage pour le diagnostic clinique de conditions ou maladies génétiques devraient participer à des épreuves de compétence (de l'anglais *proficiency testing* ou PT) externes. Notamment, le College of American Pathologists (CAP)¹¹ offre des épreuves de compétence spécialement conçues pour les laboratoires cliniques effectuant des analyses par SNG. Certaines épreuves sont spécifiques à des conditions ou maladies génétiques particulières, notamment dans le domaine de l'oncologie, alors que d'autres visent à éprouver la méthode globale de SNG et le pipeline informatique par l'analyse de données *in silico*.

¹¹ College of American Pathologists (CAP). Proficiency Testing [site Web], disponible à : <https://www.cap.org/laboratory-improvement/proficiency-testing>.

Des programmes de PT peuvent également être institués de façon indépendante entre deux laboratoires ou plus par un système d'échange d'échantillons à l'aveugle. Dans la situation où un autre laboratoire clinique ne serait pas disponible pour un système d'échange, un certain nombre d'échantillons pourraient aussi être envoyés dans un établissement central spécialisé dans le génotypage et le séquençage. Cette approche permet d'assurer la reproductibilité des résultats avec une méthode et des amorces PCR différentes. Advenant que les seuls échantillons disponibles pour des épreuves de compétence soient internes, un certain nombre d'entre eux pourrait faire l'objet d'une nouvelle analyse à l'aveugle (personnel technique, appareils de séquençage et logiciels d'analyse différents) [ACMG, 2008].

RÉFÉRENCES

- American College of Medical Genetics (ACMG). Standards and guidelines for clinical genetics laboratories. Bethesda, MD : ACMG; 2008. Disponible à : <https://www.acmg.net/PDFLibrary/Standards-Guidelines-Clinical-Molecular-Genetics.pdf>.
- American Society for Histocompatibility and Immunogenetics (ASHI). Minimum guidelines for validation of NGS for HLA. Mount Laurel, NJ : ASHI; 2015. Disponible à : <https://multibriefs.com/briefs/ashihla/NGSGuidelines.pdf>.
- Association des praticiens de génétique moléculaire (ANPGM). Recommandations ANPGM pour l'élaboration d'un compte rendu de résultats obtenus par analyse de Séquençage de Nouvelle Génération (NGS). BP-ANPGM_009. Paris, France : ANPGM; 2014. Disponible à : https://anpgm.fr/media/documents/BP-ANPGM_009-CR-NGS.pdf.
- Association des praticiens de génétique moléculaire (ANPGM). Recommandations et dossier type de validation de la méthode de séquençage par méthode de Sanger pour la recherche de mutations. BP-ANPGM_002. Paris, France : ANPGM; 2012. Disponible à : https://anpgm.fr/media/documents/BP-ANPGM_002_validation_initiale_Sanger.pdf.
- Aziz N, Zhao Q, Bry L, Driscoll DK, Funke B, Gibson JS, et al. College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. Arch Pathol Lab Med 2015;139(4):481-93.
- Beck TF, Mullikin JC, Biesecker LG. Systematic evaluation of Sanger validation of next-generation sequencing variants. Clin Chem 2016;62(4):647-54.
- Canadian Standards Association (CSA). CAN/CSA Z316.8-18 – Exigences concernant la conception, le développement et la validation d'essais conçus en laboratoire pour le dépistage, le diagnostic et la gestion des conditions cliniques. Norme nationale du Canada. Toronto, ON : CSA; 2018.
- Clinical and Laboratory Standards Institute (CLSI). MM09-A2: Nucleic acid sequence methods in diagnostic laboratory medicine; Approved guideline—Second Edition. Wayne, PA : CLSI; 2014.
- Clinical and Laboratory Standards Institute (CLSI). MM20-A: Quality management for molecular genetic testing; Approved guideline. Wayne, PA : CLSI; 2012a.
- Clinical and Laboratory Standards Institute (CLSI). MM01-A3: Molecular methods for clinical genetics and oncology testing; Approved guideline—Third Edition. Wayne, PA : CLSI; 2012b.
- Clinical and Laboratory Standards Institute (CLSI). EP12-A2: User protocol for evaluation of qualitative test performance; Approved guideline—Second Edition. Wayne, PA : CLSI; 2008.

- Clinical and Laboratory Standards Institute (CLSI). EP17-A: Protocols for determination of limits of detection and limits of quantitation; Approved guideline. Wayne, PA : CLSI; 2004.
- College of American Pathologists (CAP). Molecular Pathology Checklist. CAP Accreditation Program. Northfield, IL : CAP; 2015. Disponible à : <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=BAF04742CE945CDBA65F9BFBBFFAE4C?doi=10.1.1.700.6931&rep=rep1&type=pdf>.
- Deans Z, Watson CM, Charlton R, Ellard S, Wallis Y, Mattocks C, Abbs S. Practice guidelines for targeted next generation sequencing analysis and interpretation. Londres, Angleterre : Association for Clinical Genetic Science (ACGS); 2015. Disponible à : https://www.acgs.uk.com/media/10789/bpg_for_targeted_next_generation_sequencing_-_approved_dec_2015.pdf.
- Deans ZC, Costa JL, Cree I, Dequeker E, Edsjo A, Henderson S, et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. *Virchows Arch* 2017;470(1):5-20.
- Ellard S, Charlton R, Yau S, Gokhale D, Taylor GR, Wallace A, et al. Practice guidelines for Sanger sequencing analysis and interpretation. Londres, Angleterre : Association for Clinical Genetic Science (ACGS); 2016. Disponible à : https://www.acgs.uk.com/media/10788/acgs_sanger_sequencing_bpg_update_2016.pdf.
- Food and Drug Administration (FDA). Consideration for design, development, and analytical validation of next generation sequencing (NGS) – Based in vitro diagnostics (IVDs) intended to aid in the diagnosis of suspected germline diseases. Guidance for Stakeholders and Food and Drug Administration Staff. Silver Spring, MD : FDA; 2018. Disponible à : <https://www.fda.gov/media/99208/download>.
- Food and Drug Administration (FDA). Use of standards in FDA regulatory oversight of next generation sequencing (NGS) – Based in vitro diagnostics (IVDs) used for diagnosing germline diseases. Draft Guidance for Stakeholders and Food and Drug Administration Staff. Silver Spring, MD : FDA; 2016. Disponible à : https://downloads.regulations.gov/FDA-2016-D-1270-0002/attachment_1.pdf.
- Fox EJ, Reid-Bayliss KS, Emond MJ, Loeb LA. Accuracy of next generation sequencing platforms. *Next Gener Seq Appl* 2014;1:1000106.
- Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol* 2012;30(11):1033-6.
- Glenn TC. Field guide to next-generation DNA sequencers. *Mol Ecol Resour* 2011;11(5):759-69.

- Hume S, Nelson TN, Speevak M, McCready E, Agatep R, Feilotter H, et al. CCMG practice guideline: Laboratory guidelines for next-generation sequencing. *J Med Genet* 2019;56(12):792-800.
- Institut national d'excellence en santé et en services sociaux (INESSS). Validation analytique de tests cliniques effectués par séquençage de type Sanger en génétique humaine. Rapport rédigé par Catherine Gravel. Québec, Qc : INESSS; 2019. Disponible à : https://www.inesss.qc.ca/fileadmin/doc/INESSS/Analyse_biomedicale/INESSS_Sanger.pdf.
- Institut national du cancer (INCa). Séquençage de nouvelle génération d'un panel de gènes pour l'analyse en génétique somatique /Validation de la méthode. Boulogne-Billancourt, France : INCa; 2016. Disponible à : https://www.e-cancer.fr/content/download/148150/1859133/file/Validation%20de%20m%C3%A9thode%20NGS_FR%20FINAL%20compil%C3%A9.pdf.
- Institut national du cancer (INCa). Validation de méthodes pour la recherche de mutations en génétique somatique. Document destiné aux laboratoires de biologie moléculaire pour accompagner la démarche de validation des techniques d'analyse. Boulogne-Billancourt, France : INCa; 2014. Disponible à : <https://www.e-cancer.fr/content/download/86001/874800/file/Validation-de-methode-en-genetique-Somatique.pdf>.
- Jennings L, Van Deerlin VM, Gulley ML. Recommended principles and practices for validating clinical molecular pathology tests. *Arch Pathol Lab Med* 2009;133(5):743-55.
- Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for validation of next-generation sequencing-based oncology panels: A joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn* 2017;19(3):341-65.
- Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: A joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* 2017;19(1):4-23.
- Luthra R, Chen H, Roy-Chowdhuri S, Singh RR. Next-generation sequencing in clinical molecular diagnostics of cancer: Advantages and challenges. *Cancers (Basel)* 2015;7(4):2023-36.
- Ma X, Shao Y, Tian L, Flasch DA, Mulder HL, Edmonson MN, et al. Analysis of error profiles in deep next-generation sequencing data. *Genome Biol* 2019;20(1):50.
- Maddalena A, Bale S, Das S, Grody W, Richards S. Technical standards and guidelines: Molecular genetic testing for ultra-rare disorders. *Genet Med* 2005;7(8):571-83.
- Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet* 2016;24(10):1515.

- Mattocks CJ, Morris MA, Matthijs G, Swinnen E, Corveleyn A, Dequeker E, et al. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet* 2010;18(12):1276-88.
- Organisation internationale de normalisation (ISO). ISO 15189:2012 Laboratoires de biologie médicale — Exigences concernant la qualité et la compétence. Genève, Suisse : ISO; 2012.
- New York State Department of Health (NY State). Next Generation Sequencing (NGS) guidelines for somatic genetic variant detection. Oncology – Molecular and Cellular Tumor Markers. Albany, NY : 2018. Disponible à : <https://www.wadsworth.org/sites/default/files/WebDoc/3NextGenSeqONCOGuidelines%2012318.pdf>.
- New York State Department of Health (NY State). Guidelines for validation submissions of Next Generation Sequencing (NGS) assays under the NYS testing Category of Genetic Testing – Molecular. Albany, NY : 2015. Disponible à : http://www.wadsworth.org/sites/default/files/WebDoc/2080900015/Germline_NextGen_Validation_Guidelines.pdf.
- Peris-Vicente J, Ochoa-Aranda E, Carda-Broch S, Esteve-Romero J. Validation of human clinical genetic tests. *Mini Rev Med Chem* 2014;14(10):831-6.
- Pont-Kingdon G, Gedge F, Wooderchak-Donahue W, Schrijver I, Weck KE, Kant JA, et al. Design and analytical validation of clinical DNA sequencing assays. *Arch Pathol Lab Med* 2012;136(1):41-6.
- Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 2013;15(9):733-47.
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17(5):405-24.
- Rousseau F, Gancberg D, Schimmel H, Neumaier M, Bureau A, Mamotte C, et al. Considerations for the development of a reference method for sequencing of haploid DNA – An opinion paper on behalf of the IFCC Committee on Molecular Diagnostics, International Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2009;47(11):1343-50.
- Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: A joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn* 2018;20(1):4-27.
- Salto-Tellez M et Gonzalez de Castro D. Next-generation sequencing: A change of paradigm in molecular diagnostic validation. *J Pathol* 2014;234(1):5-10.

Van den Bulcke M, San Miguel L, Salgado R, De Quecker E, De Schutter H, Waeytens A, et al. Tests de panels de gènes par Next Generation Sequencing pour un traitement ciblé en oncologie et en hémato-oncologie – Synthèse. KCE Reports 240Bs. Bruxelles, Belgique : Centre fédéral d'expertise des soins de santé (KCE); 2015. Disponible à : https://kce.fgov.be/sites/default/files/atoms/files/KCE_240Bs_NGS_traitement_cible_Synthese.pdf.

ANNEXE A

Stratégies de repérage d'information scientifique

Bases de données bibliographiques

PubMed (NLM) Date du repérage : septembre 2016 Limites : 2000- ; anglais, français Dernière mise à jour : mai 2019	
#1	(molecular[ti] AND test[ti]) OR (molecular[ti] AND tests[ti]) OR (molecular[ti] AND testing[ti]) OR (molecular[ti] AND technique*[ti]) OR (molecular[ti] AND analysis[ti])
#2	next generation sequencing[ti] OR next-generation sequencing[ti]
#3	(genotyp*[ti] AND test[ti]) OR (genotyp*[ti] AND tests[ti]) OR (genotyp*[ti] AND testing[ti]) OR (genotyp*[ti] AND technique*[ti]) OR (genotyp*[ti] AND analysis[ti]) OR NGS[ti] OR Sanger[ti] OR TaqMan[ti] OR SnapShot[ti]
#4	standard[ti] OR standards[ti] OR validation[ti] OR validating[ti] OR verification[ti] OR norm[ti] OR norms[ti] OR accuracy[ti] OR precision[ti] OR sensitivity[ti] OR specificity[ti]
#5	Guideline[pt] OR Practice Guideline[pt] OR Guidelines as Topic[mh:noexp] OR Practice Guidelines as Topic[mh] OR Health Planning Guidelines[mh] OR Algorithms[mh] OR Clinical Conference[pt] OR Consensus[mh] OR Consensus Development Conference, NIH[pt] OR Consensus Development Conference[pt] OR Consensus Development Conferences, NIH as topic[mh] OR Consensus Development Conferences as Topic[mh] OR Critical Pathways[mh] OR Clinical Protocols[mh] OR guideline*[tiab] OR guide line*[tiab] OR guidance*[tiab] OR practical guide*[tiab] OR CPG[tiab] OR CPGs[tiab] OR algorithm*[tiab] OR best practice*[tiab] OR (best[ti] AND practice*[ti]) OR evidence base*[tiab] OR evidence report*[tiab] OR evidence syntheses*[tiab] OR research evidence*[tiab] OR practice based evidence[tiab] OR best evidence[tiab] OR clinical path[tiab] OR clinical paths[tiab] OR clinical pathway*[tiab] OR clinical protocol*[tiab] OR committee opinion*[tiab] OR consensus[tiab] OR critical pathway*[tiab] OR policy statement*[tiab] OR position statement*[tiab] OR practice parameter*[tiab] OR practice pathway*[tiab] OR practice protocol*[tiab] OR recommendation*[tiab] OR standard*[ti] OR standard of care[tiab] OR standards of care[tiab] OR standard care*[tiab] OR gold standard*[tiab] OR practice standard*[tiab] OR standard practice*[tiab] OR standard of practice*[tiab]
#6	Meta-Analysis[mh] OR Meta-Analysis[pt] OR Meta-Analysis as Topic[mh] OR Systematic Review[pt] OR Technology Assessment, Biomedical[mh] OR meta-analy*[tiab] OR metaanaly*[tiab] OR met analy*[tiab] OR metanaly*[tiab] OR meta regression*[tiab] OR metaregression*[tiab] OR meta review*[tiab] OR metareview*[tiab] OR meta synthesis[tiab] OR metasynthesis[tiab] OR overview of review*[tiab] OR overviews of review*[tiab] OR (systematic*[tiab] AND (review*[tiab] OR overview*[tiab] OR search*[tiab] OR research*[tiab])) OR (review[tw] AND (medline[tiab] OR pubmed[tiab]) AND (cinahl[tiab] OR cochrane[tiab] OR embase[tiab] OR psycinfo[tiab])) OR umbrella review*[tiab] OR technology appraisal*[tiab] OR technology assessment*[tiab] OR technology overview*[tiab] OR technology reassessment*[tiab] OR HTA[tiab] OR HTAs[tiab] OR methodological overview*[tiab] OR methodologic overview*[tiab] OR methodological review*[tiab] OR methodologic review*[tiab] OR quantitative review*[tiab] OR quantitative overview*[tiab] OR quantitative syntheses*[tiab] OR integrative review*[tiab] OR integrative overview*[tiab] OR integrative literature review*[tiab]
#7	Case Reports[pt] OR Comment[pt] OR Editorial[pt] OR Letter[pt] OR case report*[ti] OR comment*[ti] OR reply[ti] OR replies[ti] OR editorial*[ti] OR letter*[ti]
#8	(#5 OR #6) NOT #7
#9	(#1 OR #2 OR #3) AND #4 AND #8

Embase (Ovid) Date du repérage : septembre 2016 Limites : 2000- ; anglais, français Dernière mise à jour : mai 2019	
1	(next generation sequencing OR next-generation sequencing OR (genotyp* AND test) OR (genotyp* AND tests) OR (genotyp* AND testing) OR (genotyp* AND technique*) OR (genotyp* AND analysis) OR NGS OR Sanger OR TaqMan OR SnapShot OR (molecular AND test) OR (molecular AND tests) OR (molecular AND testing) OR (molecular AND technique*) OR (molecular AND analysis)).ti
2	(standard OR standards OR validation OR validating OR verification OR norm OR norms OR accuracy OR precision OR sensitivity OR specificity).ti
3	1 AND 2

EBM Reviews (Ovid) : Cochrane Database of Systematic Reviews ; Health Technology Assessment ; NHS Economic Evaluation Database Date du repérage : septembre 2016 Limites : 2000- ; anglais, français Dernière mise à jour : mai 2019	
1	(next generation sequencing OR next-generation sequencing OR (genotyp* AND test) OR (genotyp* AND tests) OR (genotyp* AND testing) OR (genotyp* AND technique*) OR (genotyp* AND analysis) OR NGS OR Sanger OR TaqMan OR SnapShot OR (molecular AND test) OR (molecular AND tests) OR (molecular AND testing) AND (molecular AND technique*) OR (molecular AND analysis)).ti

ANNEXE B

Description du principe technologique du SNG et précisions techniques

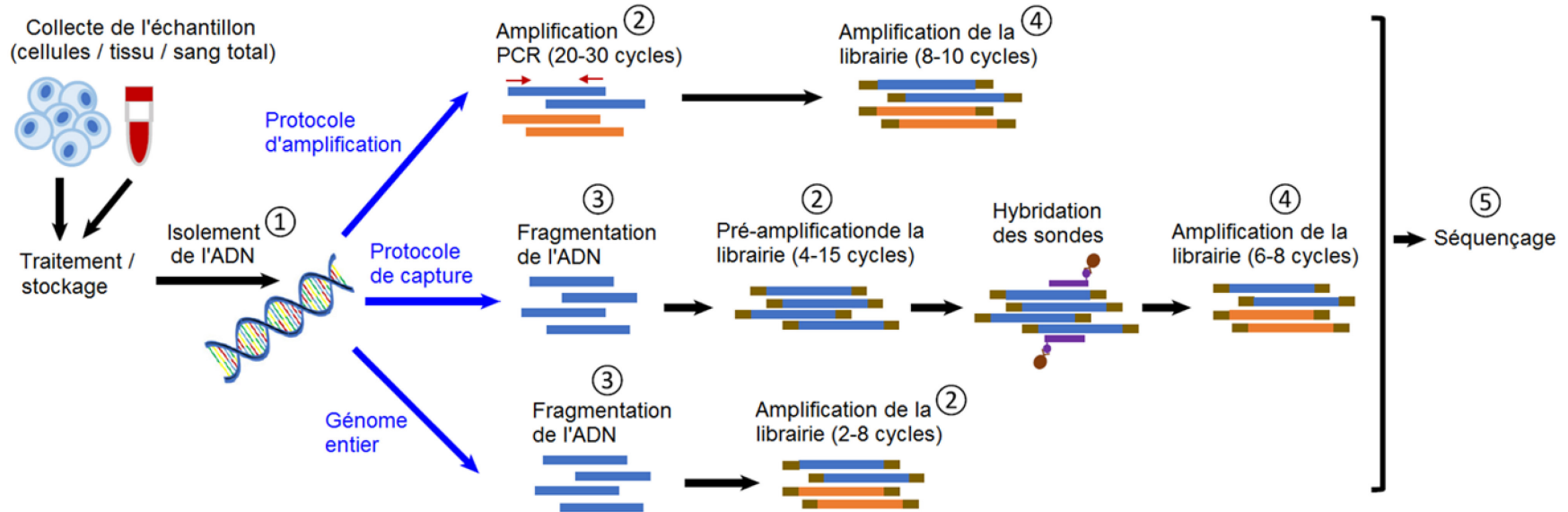
Le séquençage de nouvelle génération (SNG), aussi appelé séquençage parallèle massif (de l'anglais *massively parallel sequencing*) ou séquençage à très haut débit, fait référence aux différentes plateformes de séquençage récemment développées, qui permettent d'obtenir la séquence de plusieurs régions du génome simultanément. Cette technologie est caractérisée par un débit de lecture très élevé comparativement à la méthode de séquençage traditionnelle de type Sanger.

Plusieurs plateformes de SNG utilisant des technologies et des chimies différentes sont disponibles sur le marché. Toutefois, ces dernières partagent des caractéristiques communes. Essentiellement, le SNG est réalisé en quatre étapes principales ([figure B-1](#)) :

- 1) Préparation de la librairie (fragmentation de l'ADN génomique et ligation d'adaptateurs spécifiques à la plateforme);
- 2) Enrichissement des séquences d'intérêt (à moins que le test vise l'analyse du génome entier);
- 3) Génération clonale et séquençage;
- 4) Démultiplexage, nettoyage des séquences, alignement des séquences sur le génome, et identification et analyse des variations nucléotidiques.

De façon générale, les réactions de séquençage reposent sur un principe semblable à celui de l'approche de type Sanger, c'est-à-dire que la séquence est déterminée au cours de la synthèse cyclique d'un brin complémentaire à un des fragments d'ADN de la librairie. Selon la plateforme utilisée, la séquence est déterminée en fonction des données de luminescence, de fluorescence ou de changement de pH produites au cours de la réaction de séquençage. Les séquences ou lectures (de l'anglais *read*) obtenues pour chacun des fragments d'ADN doivent être assemblées et alignées sur une séquence de référence afin d'identifier les variations nucléotidiques [Gargis *et al.*, 2012]. Les plateformes de SNG requièrent donc des systèmes informatiques (pipelines) sophistiqués capables d'aligner les données brutes de lecture sur une séquence de référence, de reconnaître et d'annoter les variations dans la séquence, ainsi que de déterminer celles ayant une signification clinique et celles nécessitant un test de confirmation [Gargis *et al.*, 2012].

Figure B-1 Étapes générales des principales techniques de séquençage de nouvelle génération (SNG) et leurs sources d'erreurs potentielles



- ① Extraction et isolement de l'ADN (désamination de la cytosine¹, formation de 8-oxoG²)
- ② Premier enrichissement (incorporation d'erreurs précoces et biais associé à la polymérase)
- ③ Fragmentation de l'ADN (désamination de la cytosine¹, formation de 8-oxoG²)
- ④ Second enrichissement par PCR (incorporation d'erreurs précoces et biais associée à la polymérase)
- ⑤ Amplification clonale / séquençage (incorporation d'erreurs)

¹ La fixation des tissus au formaldéhyde et les traitements thermiques de l'ADN peuvent causer la désamination spontanée de la cytosine en uracile, responsable de substitutions nucléotidiques de type C:G>T:A [Ma *et al.*, 2019].

² La fragmentation ultrasonique de l'ADN peut causer l'oxydation des bases nucléotidiques, incluant la formation de 8-oxoguanine (8-oxoG) à l'origine d'un mésappariement avec l'adénine, responsable de transitions nucléotidiques de type C:G>T:A [Ma *et al.*, 2019].

Source : Traduction d'une illustration publiée par Ma et ses collaborateurs [2019].

ANNEXE C

Paramètres et mesures de qualité du SNG

Tableau C-1 Paramètres et mesures de qualité des données brutes et d'alignement des lectures des essais effectués par SNG

PARAMÈTRE	DÉFINITION
Valeur de qualité de l'appel des bases	Score permettant d'évaluer les possibilités que l'appel de la base soit erroné. Le score de qualité tend à diminuer vers la fin des lectures (diminution du signal). Une valeur de qualité médiane de l'appel des bases par cycle peut aussi être définie.
Valeur de qualité de l'alignement	Qualité de l'alignement de chaque lecture sur la séquence de référence calculée selon des algorithmes (ex. : tolérance des mésappariements, pénalité de mésappariement, pénalité de « gap », d'extension de « gap », etc.).
Pourcentage d'alignement	Fraction ou pourcentage (%) des lectures qui se sont alignées sur la séquence ou le génome de référence.
Score de « mappabilité »	Probabilité que chaque lecture s'aligne sans ambiguïté à une position donnée.
Profondeur de lecture (couverture verticale)	Nombre de fois qu'une base donnée a été séquencée ou nombre de lectures par position (parfois appelée « couverture verticale »). Elle peut aussi être définie pour une région cible (profondeur moyenne).
Couverture (couverture horizontale)	Fraction de la séquence de référence couverte par au moins une lecture (parfois appelée « couverture horizontale »). Elle est souvent exprimée en pourcentage (%) des régions d'intérêt séquencées et pour lequel la profondeur de lecture atteint ou dépasse une valeur minimale prédéfinie. Couverture moyenne : nombre moyen de lectures chevauchantes à l'intérieur d'une zone totale séquencée. Uniformité de couverture : distribution de la couverture à l'intérieur de régions spécifiques ciblées dans lesquelles l'appel de variation nucléotidique est effectué.
Duplication des lectures	Nombre de lectures avec les mêmes positions de départ. Toutes duplications ou lectures identiques peuvent représenter des copies générées lors de l'amplification et doivent être retirées afin de déterminer la profondeur de lecture réelle. Le pourcentage (%) de duplication de lecture est un indicateur de la complexité et de la qualité de la librairie.
Biais associé au brin de l'ADN	Mesure du nombre de lectures effectuées sur les brins sens et antisens. Un déséquilibre entre ces nombres peut être une source potentielle d'appel erroné de variation nucléotidique.
Biais associé au GC	Mesure de la sous-représentation des régions riches en CG. Procure une mesure de l'uniformité de couverture à l'intérieur du génome ou de régions cibles. Devrait rester constant entre les lectures. Permet d'évaluer la qualité de la préparation de l'échantillon et des étapes de capture.

Source : Tableau adapté de celui publié par le groupe Nex-StoCT [Gargis *et al.*, 2012] et modifié selon les informations supplémentaires publiées notamment par l'ACGS [Deans *et al.*, 2015], le CLSI [2014a], l'EuroGentest [Matthijs *et al.*, 2016] et l'INCa [2016].

Tableau C-2 Paramètres et mesures de qualité des données d'appel des variations nucléotidiques obtenues par SNG

PARAMÈTRE	DÉFINITION
Nombre total de variations nucléotidiques détectées	Le nombre total de variations de séquence détectées dans un échantillon clinique par rapport à une référence pour une région ciblée devrait être similaire pour une même méthode d'enrichissement et une même analyse (même panel de gènes).
Appui des lectures	Nombre de lectures qui appuient un appel de variation nucléotidique (décompte des lectures effectuées dans chacune des deux directions et équilibre de ce décompte entre les directions).
Fréquence allélique	Fraction ou pourcentage (%) de lectures alléliques contenant une variation de séquence et permettant de définir le caractère homozygote, hémizygote ou hétérozygote.
Pourcentage de polymorphisme	Pourcentage (%) des variations de séquence connues comme étant un polymorphisme. La majorité des variations nucléotidiques (> 90 %) détectées dans chaque échantillon devraient être des polymorphismes.
Valeur de qualité de l'appel d'une variation nucléotidique	Valeur p ou facteur de Bayes

Source : Tableau adapté de celui publié par le groupe Nex-StoCT [Gargis *et al.*, 2012] et modifié selon les informations supplémentaires publiées notamment par l'ACGS [Deans *et al.*, 2015], le CLSI [2014a], l'EuroGentest [Matthijs *et al.*, 2016] et l'INCa [2016].

ANNEXE D

Description des critères de performance nécessaires à la validation de la méthode ou d'un test effectué par séquençage de nouvelle génération (SNG)

Exactitude

L'objectif de déterminer l'exactitude d'une méthode est de fournir les preuves qu'elle produira les résultats escomptés. En général, l'exactitude peut être établie en déterminant les différences ou les écarts entre la valeur mesurée par la méthode et la vraie valeur (valeur de référence). Pour le SNG, la « vraie valeur » peut être définie par la séquence de référence la plus récente [Matthijs *et al.*, 2016].

Selon l'EuroGentest, la détermination de l'exactitude fait partie de la validation générale de la méthode de SNG. Elle peut entre autres être définie d'après une couverture de séquençage adéquate et le score de qualité de l'appel des bases [Matthijs *et al.*, 2016; Aziz *et al.*, 2015]. La couverture adéquate nécessaire dépend du type de variation nucléotidique présente dans la séquence et de son nombre de copies. Une plus faible couverture est nécessaire pour détecter avec précision des variations d'un seul nucléotide (SNPs) homozygotes ou hémizygotiques que pour détecter des variations hétérozygotes [Gargis *et al.*, 2012]. Le seuil de ce paramètre pour une fréquence allélique doit donc être déterminé de façon empirique lors de la validation de la méthode [Matthijs *et al.*, 2016].

D'autres paramètres de qualité, ayant notamment un effet sur la couverture, servent à l'évaluation de l'exactitude (ex. : la fraction allélique, la duplication des lectures, les biais associés au contenu en GC et au brin d'ADN).

L'exactitude doit être validée pour chaque type de variation de séquence que le laboratoire propose de détecter à l'aide de sa méthode de SNG [Matthijs *et al.*, 2016; Aziz *et al.*, 2015; Deans *et al.*, 2015; Rehm *et al.*, 2013]. L'American College of Medical Genetics (ACMG) recommande de calculer le taux moyen de résultats erronés par échantillon et de l'exprimer sous forme de nombre d'erreurs par kilobase (kb) de séquence [Rehm *et al.*, 2013].

Sensibilité

Généralement, la sensibilité correspond au rapport entre le nombre de résultats correctement identifiés comme étant positifs (vrais positifs) et l'ensemble des résultats réellement positifs (vrais positifs et faux négatifs).

En SNG, la sensibilité d'un test est définie par le taux de faux négatifs. Les faux négatifs peuvent provenir de plusieurs sources, notamment d'une couverture insuffisante d'une région d'intérêt, d'un défaut de l'appel d'un variant en raison de l'utilisation de paramètres de qualité inférieure, ou d'une distribution des lectures des brins sens et antisens ne répondant pas aux critères de contrôle de qualité [INCa, 2016; Aziz *et al.*, 2015].

Le taux de faux négatifs et la sensibilité d'un test en particulier peuvent également être évalués lors d'une étude de concordance (voir la [section 6.3](#) du présent document). L'American College of Medical Genetics (ACMG) recommande de calculer le taux moyen de faux négatifs par échantillon et de l'exprimer sous forme de nombre de faux positifs par kilobase (kb) de séquence [Rehm *et al.*, 2013].

Spécificité

Généralement, la spécificité correspond au rapport entre le nombre de résultats correctement identifiés comme étant négatifs (vrais négatifs) et l'ensemble des résultats réellement négatifs (vrais négatifs et faux positifs).

La spécificité est généralement calculée en utilisant des échantillons négatifs afin de déterminer la fraction correctement identifiée comme ne comportant pas de variations de séquence. Ce critère est généralement déterminé lors de la validation de la méthode, par le calcul de la spécificité de la région analysée, par exemple en déterminant la fraction de variations nucléotidiques correspondant à un appel de base incorrect (taux de faux positifs) [Aziz *et al.*, 2015]. Outre une erreur de l'appel de la base, un faux positif peut être attribuable à une erreur d'alignement de la séquence (présence de pseudogènes ou de séquences homologues) [Aziz *et al.*, 2015; Pont-Kingdon *et al.*, 2012]. L'American College of Medical Genetics (ACMG) recommande de calculer le taux moyen de faux positifs par échantillon et de l'exprimer sous forme de nombre de faux positifs par kilobase (kb) de séquence [Rehm *et al.*, 2013].

Les faux positifs peuvent être vérifiés et confirmés par l'utilisation de méthodes alternatives (ex. : séquençage Sanger) [Aziz *et al.*, 2015], notamment, lors d'une étude de concordance.

La spécificité acceptable d'un test varie selon l'application clinique pour laquelle il est prévu (dépistage, confirmation, diagnostic) et selon la mise en place (ou non) d'un processus de confirmation des variations nucléotidiques détectées avant de les rapporter (ex. : confirmation par séquençage Sanger ou par une méthode orthogonale) [Hume *et al.*, 2019].

Précision

Selon l'EuroGentest, la détermination de la précision fait partie de la validation générale de la méthode de SNG [Matthijs *et al.*, 2016]. Pour les tests qualitatifs comme le séquençage de l'ADN, la précision fait référence à la capacité de la méthode à reproduire les mêmes résultats pour un même échantillon, lorsqu'elle est employée à répétition. Afin d'évaluer la robustesse et la fiabilité de la méthode, des études de précision intra-essai (répétabilité) et inter-essai (fidélité intermédiaire) sont effectuées. Une concordance complète des résultats n'est pas possible pour les technologies de SNG, toutefois le laboratoire doit établir des intervalles de concordance jugés suffisants afin de définir les paramètres pour une répétabilité, une fidélité et une reproductibilité acceptables [Rehm *et al.*, 2013].

Répétabilité (précision intra-essai)

La précision intra-essai est évaluée par l'analyse simultanée de quelques échantillons en triplicata (au minimum) lors du même essai, ou lors d'essais différents effectués à partir des mêmes conditions d'analyse (ex. : à l'aide du même appareillage, du même lot de réactifs, du même opérateur) [CLSI, 2014].

Fidélité intermédiaire (précision inter-essai)

La précision inter-essai est évaluée à l'aide des mêmes échantillons testés lors d'essais indépendants effectués à des jours différents en faisant varier certains paramètres propres à l'activité du laboratoire (ex. : lots de réactifs différents, opérateurs différents) [CLSI, 2014; Pont-Kingdon *et al.*, 2012].

Reproductibilité (précision inter-laboratoire)

La précision inter-laboratoire est évaluée à l'aide des mêmes échantillons testés lors d'essais indépendants effectués par des laboratoires différents.

Limite de détection

La limite de détection (de l'anglais *limit of detection* ou LOD) correspond à la plus faible fraction allélique pouvant être détectée dans un échantillon.

La LOD ne s'applique généralement pas aux analyses de séquençage d'ADN utilisées pour la détection de mutations germinales associées à un désordre héréditaire. Ce paramètre doit toutefois être déterminé dans les cas où l'analyse est utilisée afin de détecter de faibles niveaux d'ADN muté parmi une population d'ADN normal. La LOD est applicable notamment dans les cas de variations de séquences minoritaires tels l'hétéroplasmie mitochondriale, le mosaïcisme, les mutations somatiques en oncologie [Ellard *et al.*, 2016; INCa, 2014; Pont-Kingdon *et al.*, 2012; Mattocks *et al.*, 2010] et la détection prénatale des aneuploïdies.

La limite de détection correspond au seuil de positivité, c'est-à-dire au seuil au-delà duquel le séquençage est capable de détecter la présence d'une variation nucléotidique par rapport à la séquence de référence. Elle correspond à la plus faible proportion d'allèles mutés pouvant être distinguée du bruit de fond. La méthode de séquençage doit alors permettre la détection d'une variation nucléotidique présente en proportion supérieure au seuil établi par le laboratoire. Ce seuil peut être défini expérimentalement ou selon la revue de la littérature. Il peut être différent en fonction des paramètres étudiés [ANPGM, 2012].

Pour les désordres mitochondriaux et les cancers, des variations peuvent être détectées à des fréquences alléliques de 1 % à 100 %. C'est pourquoi ces circonstances requièrent qu'une limite de détection soit établie pour la fréquence allélique [Hume *et al.*, 2019]. De façon générale, la LOD minimale à atteindre en génétique somatique est de 10 % d'allèles mutés pour les cibles d'intérêt clinique validées [INCa, 2016]. Des niveaux de mutation inférieurs à la LOD produiront des résultats faussement négatifs et affecteront la sensibilité du test [INCa, 2014; Pont-Kingdon *et al.*, 2012].

Lorsque cela est nécessaire, la LOD d'un test par séquençage peut être déterminée par une étude de dilution décroissante. L'ADN d'un échantillon comportant une variation pathogénique (hétérozygote ou homozygote) connue peut être dilué à différentes concentrations dans un échantillon d'ADN normal [INCa, 2014; Pont-Kingdon *et al.*, 2012].

Limite de quantification

Selon le CLSI, la limite de quantification (en anglais *limit of quantification* ou LOQ) est la plus petite quantité de mesurande, dans un échantillon, qui peut être déterminée quantitativement avec une précision et une exactitude acceptables et dans des conditions expérimentales prédéfinies [CLSI, 2004]. Elle se mesure au cours du même devis expérimental que la LOD, en utilisant plusieurs courbes de dilution d'échantillons différents et en déterminant la variabilité de la mesure sur la courbe de dilution à différentes concentrations de mesurande. Dans le cas du séquençage, il s'agit de la plus petite fraction allélique, à un ou plusieurs nucléotides donnés, pouvant être mesurée avec une exactitude et une précision acceptables et prédéfinies, et ce, en fonction du besoin clinique.

Intervalle de mesure

En séquençage, l'intervalle de mesure correspond à la région du génome pour laquelle une séquence de qualité acceptable peut être obtenue.

Pour le SNG, ce paramètre peut comporter la définition de la région du gène interrogée, les régions d'intérêt ou la distribution des mutations si elles sont connues, le détail de tous les résultats envisageables pouvant être détectés par la méthode ou la technique, notamment le statut mutationnel, homozygote ou hétérozygote, ou toute autre information pertinente concernant la séquence du gène d'intérêt [CAP, 2015; CLSI, 2014; Pont-Kingdon *et al.*, 2012]. Dans le cas du SNG, ce paramètre est plutôt défini lors de la planification préanalytique (voir la [section 6](#) du présent document) [Mattocks *et al.*, 2010].

Intervalle de référence

Pour le séquençage, ce paramètre représente la séquence de référence utilisée pour l'interprétation des résultats [CLSI, 2014; Pont-Kingdon *et al.*, 2012]. Elle doit être bien caractérisée et permettre de distinguer les variations normales (ex. : polymorphisme) de celles qui sont anormales. Dans le cas du SNG, ce paramètre est également défini lors de la planification préanalytique (voir la [section 6.1](#) du présent document pour plus de détails) [Mattocks *et al.*, 2010].

Étude de concordance

En plus de la validation analytique des paramètres et critères de qualité, qui peut être en partie effectuée par l'utilisation de matériel de référence, une étude de concordance doit être réalisée à l'aide de spécimens du même type que ceux qui seront acceptés pour l'analyse, et ce, pour déterminer la sensibilité, la spécificité et l'exactitude d'un nouveau test. Pour ce faire, des échantillons cliniques caractérisés contenant des variations de

séquence connues sont analysés (voir la [section 6.3](#) du présent document pour plus d'informations concernant les types d'échantillons témoins) [Gargis *et al.*, 2012].

Autrement, les variations détectées dans un échantillon particulier peuvent être vérifiées par une autre méthode. Plusieurs techniques peuvent être utilisées comme méthodes de comparaison pour valider un test par SNG. Parmi celles-ci figurent notamment :

- le séquençage Sanger;
- les essais de génotypage de variations nucléotidiques (SNP);
- les méthodes de SNG orthogonales;
- l'utilisation d'autres techniques moléculaires telles l'amplification multiplex de sondes dépendante d'une ligation (de l'anglais *multiplex ligation-dependent probe amplification* ou MLPA) et l'hybridation génomique comparative (*comparative genomic hybridization* ou CGH) pour la confirmation des délétions et des duplications [Deans *et al.*, 2015].

Les études de concordance doivent être effectuées à l'aveugle (sans que l'opérateur connaisse les résultats de la méthode de comparaison) [Deans *et al.*, 2015; Pont-Kingdon *et al.*, 2012; Mattocks *et al.*, 2010].

ANNEXE E

Recommandations d'organismes professionnels et réglementaires concernant les critères de performance à évaluer lors de la validation analytique d'un test clinique effectué par SNG ou dans le contexte des analyses génétiques et moléculaires

ORGANISME [AUTEUR(S), ANNÉE]	SENSIBILITÉ	SPÉCIFICITÉ	EXACTITUDE	PRÉCISION		LOD*	INTERVALLE DE RÉFÉRENCE	INTERVALLE DE MESURE	ÉTUDE DE CONCORDANCE	CHAMPS D'ACTIVITÉ
				INTRA- ESSAI	INTER- ESSAI					
ACGS [Deans <i>et al.</i> , 2015]	++	n. d.	n. d.	++	++	n. d.	n. d.	n. d.	n. d.	Génétique héréditaire
ACMG [Rehm <i>et al.</i> , 2013]	++	++	++	n. d.	++	n. d.	n. d.	n. d.	++	Génétique héréditaire
ASHI [ASHI, 2015]	n. d.	n. d.	++	++	++	n. d.	n. d.	n. d.	++	Histocompatibilité Immunogénétique
CAP [Aziz <i>et al.</i> , 2015]	++	++	++	++	++	++	n. d.	n. d.	n. d.	Laboratoires médicaux
CLSI [CLSI, 2014]	++	++	++	++	++	++	++	++	n. d.	Laboratoires médicaux
EuroGentest [Matthijs <i>et al.</i> , 2016]	++	++	++	++		n. d.	++	++	n. d.	Génétique héréditaire
FDA [FDA, 2018]	++	++	++	++	++	++	n. d.	n. d.	++	Génétique héréditaire
INCa [INCa, 2016]	++	++	n. d.	+	++	++	n. d.	n. d.	++	Génétique somatique
État de New York [NY State, 2015]	++	++	n. d.	n. d.	++	++	n. d.	n. d.	++	Laboratoires médicaux
PSGBI [Salto-Tellez et Gonzalez de Castro, 2014]	++	++	++	n. d.	n. d.	n. d.	n. d.	n. d.	n. d.	Génétique somatique

Abréviations : ACGS : Association for Clinical Genetic Science (Royaume-Uni); ACMG : American College of Medical Genetics and Genomics (États-Unis); ASHI : American Society for Histocompatibility and Immunogenetics (États-Unis); CAP : College of American Pathologists (États-Unis); CCMG : Canadian College of Medical Geneticists (Canada); CLSI : Clinical & Laboratory Standards Institute (États-Unis); FDA : Food and Drug Administration (États-Unis); INCa : Institut national du cancer (France); LOD : limite de détection (de l'anglais *limit of detection*); n. d. : donnée non disponible; PSGBI : Pathological Society of Great Britain and Ireland.

* Applicable dans les cas de variations de séquences minoritaires tels l'hétéroplasmie mitochondriale, le mosaïcisme et les échantillons tumoraux (oncologie) [Mattocks *et al.*, 2010].

++ Critère recommandé.

+ Critère suggéré.

ANNEXE F

Validation analytique d'un test clinique effectué par séquençage de nouvelle génération (SNG) : recommandations de bonnes pratiques

RAPPEL

Les éléments à définir ou à évaluer ne constituent pas des normes mais bien des recommandations de bonnes pratiques. L'INESSS est conscient qu'il peut exister différents types de contraintes pouvant empêcher un laboratoire, pour une analyse donnée, de réaliser toutes les mesures proposées ici.

Tableau F-1 Éléments et paramètres à définir ou à évaluer lors de la planification préanalytique d'un test effectué par SNG

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
1. Objectif de l'analyse (section 5.1)			
1.1 Préciser s'il s'agit : - d'un test de dépistage, de confirmation ou de diagnostic; - d'une recherche de variations nucléotidiques constitutives (germinales) ou acquises (somatiques); - d'une recherche de variations(s) pathogénique(s) connue(s) ou d'une caractérisation de région(s) définie(s) du génome (ex. : gène ou panel de gènes associés à une condition particulière).	S. O.	S. O.	

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
2. Identification des séquences de référence (section 6.1)			
2.1 Indiquer le numéro d'identification des séquences de référence. Utiliser une référence reconnue (ex. : LRG, RefSeq ou GeneBank^{MC}) ou la version du génome utilisée (ex. : GRCh38.p9).	S. O.	S. O.	
2.2 Fournir ou indiquer les éléments suivants : régions spécifiques du génome qui seront analysées par le test, y compris leur localisation sur la séquence de référence (ex. : gènes complets ou partiels, incluant ou non les régions promotrices, les introns, etc.); - informations rapportées dans la littérature scientifique concernant les régions interrogées par le test, et ce, en lien avec la pathologie ou la condition médicale à l'étude et permettant de justifier : - le choix des régions analysées (ex. : gènes ou exons d'intérêt, sites de variations pathogéniques connues, <i>hotspot</i>, etc.), - l'exclusion de régions particulières, le cas échéant (ex. : pseudogènes, régions homologues, régions riches en CG, etc.).	S. O.	S. O.	

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
3. Conception, choix et validation des amorces (section 6.2)			
3.1 Indiquer le type de chimie et le nom de la plateforme de SNG utilisées (ex. : séquençage par synthèse, par ligation, protocole en PCR, en capture, etc.).	s. o.	s. o.	
3.1 Spécifier la provenance des amorces utilisées (ex. : panel ou trousse commerciale, panel sur mesure).	s. o.	s. o.	
3.2 Dans le cas d'un test conçu en laboratoire (TCL) utilisant une méthode de séquençage basé sur une amplification PCR , indiquer : - la stratégie adoptée pour minimiser la perte allélique (<i>allelic drop-out</i>); - la limite de tolérance fixée concernant la présence de polymorphismes au site de liaison des amorces (fréquence acceptable et positionnement en 3' des polymorphismes) afin de limiter la perte allélique; - la stratégie adoptée pour éviter l'amplification de régions homologues ou de pseudogènes, le cas échéant.	Exemple : Séquençage de deux amplicons différents (paires d'amorces différentes) pour chaque locus analysé (méthodes ciblées) < 1 % de fréquence allélique dans n'importe quelle population dans les 5 dernières bases de l'amorce		
4. Détermination du matériel de référence et des échantillons cliniques témoins (sections 6.3 et 6.4)			
4.1 Fournir la liste des types de spécimens cliniques qui pourront être soumis au laboratoire aux fins d'analyse.	Exemple : sang (leucocytes), frottis buccal, cellules en culture, etc.	Exemple : tissus tumoraux frais, tissus FFPE, moelle osseuse, etc.	

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
4.2 Indiquer le matériel de référence (de préférence certifié) utilisé pour la validation analytique du test, y compris : - la provenance (ex. : HapMap, NIST, laboratoire certifié, population particulière); - la nature (ex. : ADN synthétique, lignée cellulaire, échantillons cliniques caractérisés); - le nombre et les types de variations nucléotidiques que comporte cette référence.	Exemple : ≥ 1 échantillon de référence comportant ≥ 300 variations nucléotidiques uniques, représentant tous les types que le test est censé détecter (ex. : SNV, MNV, CNV, insertion, délétion, etc.)		
4.3 Indiquer les échantillons cliniques utilisés pour l'étude de validation analytique du test (étude de concordance), y compris : - le type (représentatif des échantillons qui seront analysés); - la variété (représentative de la population ciblée et, dans la mesure du possible, de tous les résultats possibles); - la méthode avec laquelle ils ont été testés pour les variations pathogéniques d'intérêt (ex. : séquençage Sanger, méthode orthogonale).	s. o.		
4.4 Spécifier quels échantillons serviront de témoins lors des essais (contrôles positifs et négatifs de qualité interne).	Contrôle(s) négatif(s) : ≥ 1 par protocole Contrôle(s) positif(s) : ≥ 1 hétérozygote et ≥ 1 homozygote (si applicable) pour chacun des marqueurs s'il s'agit d'un test ciblant des variations pathogéniques spécifiques		

Tableau F-2 Éléments et paramètres à définir ou à évaluer lors de la [validation analytique](#) d'un test effectué par SNG

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
1. Pipeline informatique clinique (section 7.1)			
1.1 Décrire le pipeline informatique , y compris les infrastructures informatiques, les mécanismes de protection des données cliniques, les logiciels utilisés et leur version, etc.	s. o.		
2. Évaluation de la qualité des données obtenues par SNG (section 7.2)			
2.1 Indiquer les valeurs minimales requises pour juger de la qualité d'un essai de séquençage et de l'appel d'une variation nucléotidique en fonction des paramètres suivants :	s. o.	s. o.	
▪ Appel des bases (valeur Phred ou autre VQ)	Exemple : ≥ 30 pour chaque base appelée		
▪ Alignement des bases et des lectures (<i>reads</i>) sur la référence	Exemple : ≥ 99 % des bases ≥ 99 % des lectures	Exemple : > 80 % (<i>clusters passing filter</i>)	
▪ Couverture moyenne et uniformité de couverture (couverture horizontale)	s. o.	s. o.	
▪ Profondeur de lecture (couverture verticale) pour l'appel d'un variant	Exemple : ≥ 30X	Exemple : ≥ 250X par cible ≥ 1 000X pour les tissus hétérogènes de faible cellularité tumorale	

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
▪ Fréquence allélique pour l'appel d'une variation nucléotidique	20 % des lectures pour appeler l'hétérozygotie	En fonction de la limite de détection du test	
2.2 Indiquer tout autre critère de qualité pertinent pour l'analyse en fonction de la plateforme de SNG et du pipeline informatique utilisé, du type d'analyse, de la limite de détection, ou autre élément.	s. o.	s. o.	
3. Critères de performance nécessaires à la validation d'un test effectué par SNG (sections 7.4 à 7.6)			
3.1 EXACTITUDE (%) Degré de concordance entre la séquence nucléotidique déterminée par le test et la séquence de référence $\frac{VP + VN}{(VP + FP + VN + FN)} \times 100$	Note : Pour les analyses qualitatives comme le séquençage, l' exactitude , la sensibilité et la spécificité sont des critères interdépendants. Les données recueillies afin d'évaluer la spécificité et la sensibilité sont utilisées pour déterminer l'exactitude. Les calculs de l'exactitude, de la sensibilité et de la spécificité d'un test de SNG s'effectuent selon l'analyse d'un nombre de variations nucléotidiques et d' échantillons cliniques caractérisés . Par exemple, utiliser : 1) ≥ 1 référence commerciale (ex. : HapMap) avec tous les types de variations nucléotidiques détectables par le test (ex. : SNV, CNV, etc.);		
3.2 SENSIBILITÉ Probabilité qu'un test détecte une variation de séquence lorsqu'elle est présente $\frac{VP}{(VP + FN)} \times 100$			
3.3 SPÉCIFICITÉ Probabilité qu'un test ne détecte pas de variation de séquence lorsqu'aucune n'est présente $\frac{VN}{(VN + FP)}$			

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
	2) ≥ 30 échantillons cliniques caractérisés, dont ≥ 15 comportant les variations pathogéniques les plus communes associées à la condition clinique à l'étude; 3) une variété de spécimens cliniques du même type que ceux qui seront analysés (ex. : sang, moelle).		
3.4 RÉPÉTABILITÉ (précision intra-essai) Degré avec lequel la même séquence est obtenue lorsqu'un même échantillon est séquencé à de multiples reprises sous les mêmes conditions	Exemple : ≥ 3 échantillons analysés à trois reprises dans les mêmes conditions de séquençage (ex. : même opérateur, utilisant le même lot de réactifs et le même instrument, le même jour)		
3.5 FIDÉLITÉ INTERMÉDIAIRE (précision inter-essai) Degré avec lequel la même séquence est obtenue lorsqu'un même échantillon est séquencé à de multiples reprises sous des conditions différentes Indiquer les variables testées (ex. : opérateurs, lots de réactifs, séquenceurs, etc.).	Exemple : ≥ 3 échantillons analysés à trois reprises en faisant varier les facteurs les plus significatifs par rapport à l'activité du laboratoire (ex. : différents opérateurs, utilisant différents lots de réactifs et sur plus d'un instrument, à des jours différents)		

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
3.6 LIMITE DE DÉTECTION (LOD) Fraction allélique minimum détectable dans un échantillon donné. La LOD correspond au seuil de positivité au-delà duquel le test de séquençage peut détecter la présence d'une variation nucléotidique comparativement à la séquence de référence. <i>Applicable dans les cas de variations de séquences minoritaires tels l'hétéroplasmie mitochondriale, le mosaïcisme, les échantillons tumoraux et l'ADN fœtal ou tumoral libre circulant.</i>	<u>Optionnel</u> À déterminer, si jugé cliniquement pertinent, selon l'application prévue du test	<u>Obligatoire</u> À déterminer en fonction de la limite souhaitée et de la profondeur de lecture du SNG Exemple : ≥ 1 échantillon à fraction allélique connue, dilué séquentiellement pour atteindre un seuil cliniquement pertinent	
3.7 LIMITE DE QUANTIFICATION (LOQ) Fraction allélique minimum mesurable avec une précision et une exactitude prédéfinies. <i>Applicable dans les cas de variations de séquences minoritaires, tels l'hétéroplasmie mitochondriale, le mosaïcisme, les échantillons tumoraux et l'ADN libre circulant fœtal ou tumoral, ou de mesures quantitatives.</i>	<u>Optionnel</u> À déterminer, si jugé cliniquement pertinent, selon l'application prévue du test.	<u>Obligatoire</u> Utiliser les mêmes courbes de dilution que celles utilisées pour la détermination de la LOD. Déterminer la variabilité de la mesure sur la courbe de dilution à différentes proportions de variations pathogéniques.	

ÉLÉMENTS ET PARAMÈTRES À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
3.8 INTERVALLE DE MESURE Région du génome pour laquelle une séquence de qualité acceptable peut être obtenue par le test.	Peut être déterminé par l'analyse des résultats de séquençage obtenus pour déterminer la sensibilité, la spécificité, la précision et la LOD.		

Tableau F-3 Éléments et paramètres à fournir, à définir ou à évaluer lors des étapes postanalytiques d'un test effectué par SNG

ÉLÉMENTS ET PARAMÈTRES À FOURNIR, À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
1. Interprétation des variations nucléotidiques et communication des données de séquence (sections 8.1 à 8.3)			
1.1 Fournir un exemple de rapport destiné au médecin demandeur. Un rapport devrait contenir minimalement les éléments suivants : a) Identification du spécimen-patient : nom, date de naissance, date de prélèvement, nature du spécimen (sang, moelle, tissu tumoral, etc.), numéro assigné au spécimen ou au patient, numéro assigné au pédigrée familial ou à d'autres membres de la famille; b) Indication du test : condition, maladie ou marqueurs testés, objectif de la demande (dépistage, confirmation diagnostique, statut de porteur, etc.) et raison sous-jacente à la	S. O.	S. O.	

ÉLÉMENTS ET PARAMÈTRES À FOURNIR, À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
demande (retard de développement, histoire familiale, cancer actif); c) Nom du prescripteur; d) Plateforme de SNG utilisée : nom et description brève de la technologie; e) Nom de l'analyse effectuée : nom du panel ou de la trousse commerciale ou de l'analyse développée localement; f) Objectif du test : recherche de variations pathogéniques connues, caractérisation de régions définie du génome (ex. : promoteur, gènes, exons, introns testés, etc.); g) Sensibilité estimée du test; h) Informations sur la qualité de l'appel d'une variation nucléotidique (variable selon les logiciels et les algorithmes utilisés par le laboratoire); i) Interprétation des résultats selon la nomenclature reconnue pour la classification des variations nucléotidiques détectées; j) Nom et coordonnées du laboratoire; k) Date et nombre de pages du rapport; l) Nom du professionnel certifié (médecin de laboratoire ou biochimiste clinique) ayant interprété et signé le rapport; m) Limites du test ou de la méthode de séquençage utilisée;			

ÉLÉMENTS ET PARAMÈTRES À FOURNIR, À DÉFINIR OU À ÉVALUER (recommandations de bonnes pratiques)	DEVIS EXPÉRIMENTAL PROPOSÉ SELON LA NATURE DES VARIATIONS NUCLÉOTIDIQUES RECHERCHÉES		INFORMATIONS OU RÉSULTATS DE VALIDATION Des informations supplémentaires peuvent être fournies en annexe de ce formulaire.
	VARIATIONS CONSTITUTIVES	VARIATIONS ACQUISES	
n) Avis de non-responsabilité aux endroits appropriés mentionnant la possibilité d'erreurs hors du contrôle du laboratoire.			
1.2 Fournir les références scientifiques lorsque des données publiées peuvent influencer l'interprétation ou le calcul du risque ou lorsque celles-ci sont conflictuelles.	S. O.	S. O.	
1.3 Décrire le processus d'analyse des données, de l'interprétation et de la validation des résultats, y compris le rôle des différents intervenants.	S. O.	S. O.	
1. Épreuves de compétence et assurance qualité (section 8.3)			
1.1 Fournir la liste et les modalités des épreuves de compétence externes de SNG auxquelles le laboratoire participe.	Épreuves de compétence de séquençage offertes par des organismes indépendants (ex. : CAP) et/ou échanges de spécimens cliniques deux fois par année avec d'autres laboratoires.		
1.2 Fournir les preuves de compétence en SNG du laboratoire.	Exemple : Certification ou preuve de passation d'une épreuve de compétence délivrée par un organisme indépendant, résultats des analyses effectuées en parallèle avec un autre laboratoire (étude de concordance), etc.		



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

