



Approche de surveillance de la population afin de détecter les nouveaux agrégats de cas de COVID-19 et y répondre

Erin E Rees^{1*}, Rachel Rodin², Nicholas H Ogden¹

Résumé

Contexte : Pour maintenir le contrôle de l'épidémie de maladie à coronavirus 2019 (COVID-19) lorsque les mesures de confinement seront levées, il sera essentiel de renforcer les mesures de santé publique alternatives. En ce qui concerne la surveillance, il sera nécessaire de détecter rapidement une grande proportion de tous nouveaux cas afin de pouvoir les isoler, et retrouver et mettre en quarantaine les personnes qui y ont été exposées. Nous présentons ici une approche mathématique qui peut être utilisée pour déterminer combien d'échantillons doivent être recueillis par unité de surface et par unité de temps pour détecter de nouveaux agrégats de cas de COVID-19 à un stade suffisamment précoce pour contrôler une éclosion.

Méthodes : Nous présentons une méthode de détermination de la taille de l'échantillon qui utilise une approche de pondération relative. Compte tenu du fait que les résultats du test de diagnostic de la COVID-19 provenant de sous-populations ont contribué à détecter la maladie à un niveau de prévalence seuil pour contrôler l'éclosion, il s'agissait de 1) déterminer si le nombre prévu d'échantillons hebdomadaires fournis par la surveillance actuelle des infections virales respiratoires fondée sur les soins de santé peut fournir une taille d'échantillon déjà adéquate pour détecter de nouveaux agrégats de cas de COVID-19 et, dans le cas contraire, 2) de déterminer combien d'échantillons hebdomadaires supplémentaires étaient nécessaires à partir d'un échantillonnage volontaire.

Résultats : Lors d'une démonstration de notre méthode à une fréquence hebdomadaire et à l'échelle des provinces et territoires canadiens, nous avons constaté que seuls les provinces et les territoires les plus peuplés disposaient d'un nombre suffisant de dépistages provenant des visites médicales en raison de maladies respiratoires pour détecter la COVID-19 à notre niveau de prévalence cible — qui doit être suffisamment élevé pour identifier et contrôler les nouveaux agrégats de cas. En outre, la détection de la COVID-19 est plus efficace (moins d'échantillons requis) lorsque la surveillance se concentre sur l'exigence de tests de dépistage de patients symptomatiques par les services des soins de santé. Dans les populations volontaires : plus les taux de contact sont élevés, plus le niveau de prévalence attendu est élevé, et moins il faut d'échantillons pour détecter la COVID-19 à un seuil prédéterminé.

Conclusion : Cette étude présente une stratégie de surveillance ciblée, combinant des échantillons de surveillance passive et active, afin de déterminer le nombre d'échantillons à recueillir par unité de surface et par unité de temps pour détecter de nouveaux agrégats de cas de COVID-19. L'objectif de cette stratégie est de permettre une détection suffisamment précoce pour contrôler une éclosion.

Citation proposée : Rees EE, Rodin R, Ogden NH. Approche de surveillance de la population afin de détecter les nouveaux agrégats de cas de COVID-19 et y répondre. *Relevé des maladies transmissibles au Canada* 2021;47(5/6):266–74. <https://doi.org/10.14745/ccdr.v47i56a01f>

Mots-clés : surveillance, détection, COVID-19, éclosion, approche mathématique

Cette oeuvre est mise à la disposition selon les termes de la licence internationale Creative Commons Attribution 4.0



Affiliations

¹ Division des sciences des risques pour la santé publique, Laboratoire national de microbiologie, Agence de la santé publique du Canada, Saint-Hyacinthe, QC et Guelph, ON

² Direction générale de la prévention et du contrôle des maladies infectieuses, Agence de la santé publique du Canada, Ottawa, ON

*Correspondance :

erin.rees@canada.ca



Introduction

Comme de nombreux pays dans le monde, le Canada a mis en place des mesures de confinement pour contrôler la transmission du virus à l'origine de la maladie à coronavirus 2019 (COVID-19) : coronavirus du syndrome respiratoire aigu sévère 2 (SRAS-CoV-2). Les mesures de confinement les plus courantes comprennent les restrictions de déplacement et la fermeture des lieux de rassemblement social tels que les restaurants, les bars et autres lieux de divertissement en intérieur. La décision de lever, de réduire ou d'arrêter les mesures de confinement tient compte de plusieurs critères, dont des considérations sociales, économiques et sanitaires, et les décisions sur l'étendue et la durée du confinement sont contrôlées aux niveaux fédéral, municipal et provincial et territorial. Au niveau le plus simpliste, les mesures de confinement peuvent être assouplies à un niveau de prévalence défini; une stratégie utilisée par l'Allemagne lors de son processus de levée des mesures de confinement après la première vague de cas de COVID-19 (1). Pour maintenir le contrôle de l'épidémie lorsque les mesures de confinement sont levées, il est essentiel de renforcer les mesures alternatives de santé publique (recherche des contacts, mise en quarantaine). Plus précisément, nous devons détecter rapidement une grande proportion de nouveaux cas afin de pouvoir les isoler, et les personnes qui ont été exposées ou en contact avec ces cas doivent être retrouvées et mises en quarantaine. Si les capacités de dépistage et de recherche des contacts sont insuffisantes, il est probable que l'épidémie resurgisse et que les services de santé soient dépassés (2,3).

La capacité de détecter une maladie au sein d'une population dépend du type de stratégie de surveillance et du nombre requis d'échantillons à tester. Dans les grandes populations (i.e. supérieures à 1 000), une approche courante suppose un échantillonnage aléatoire parmi les individus qui ont un risque égal d'obtenir un résultat positif au dépistage de la maladie (4). Toutefois, si l'on dispose d'informations sur les caractéristiques qui contribuent à la probabilité d'un dépistage positif, il est alors possible d'utiliser une approche ciblée pour déterminer de façon optimale la taille de l'échantillon, en pondérant les échantillons en fonction de leur capacité de détection compte tenu de leurs caractéristiques (5).

Au Canada, il existe actuellement deux stratégies principales de collecte d'échantillons pour la surveillance de la COVID-19 : 1) les visites médicales et les admissions à l'hôpital en raison d'une maladie respiratoire (demande de dépistage de patients symptomatiques par les services de soins de santé); et 2) les populations à risque, telles que les travailleurs essentiels, qui craignent d'avoir été exposés à l'infection (6–8). Cependant, ces méthodes peuvent ne pas permettre d'obtenir une taille d'échantillon suffisamment importante pour détecter de nouveaux agrégats de cas de transmission à un moment suffisamment précoce (i.e. lorsque la prévalence de l'infection dans la communauté est faible) pour veiller à ce que la capacité de santé publique soit en mesure de retrouver et mettre en

quarantaine les contacts afin de contrôler la transmission. Pour obtenir une taille d'échantillon suffisante, il peut être nécessaire de recourir à une stratégie d'échantillonnage faisant appel à des volontaires. Cela permettrait probablement de détecter davantage de cas asymptomatiques que lors de l'échantillonnage des personnes qui consultent les services de santé; néanmoins, l'intérêt d'inclure l'échantillonnage de la population volontaire serait double : premièrement, améliorer l'alerte précoce en effectuant des tests plus largement dans la communauté et augmenter ainsi la probabilité de détecter de nouveaux agrégats de cas; et deuxièmement, déclencher une réponse de santé publique à un niveau déterminé de prévalence dans la population à partir duquel il est possible de contrôler l'épidémie sans devoir remettre en place des mesures de confinement généralisées.

Les stratégies de surveillance ciblée peuvent être utilisées pour échantillonner efficacement une population, qui contient des sous-populations ayant des probabilités différentes d'être infectées, lorsque l'objectif est la détection précoce de la maladie à un niveau de prévalence donné (9,10). Cette approche nécessite une évaluation des échantillons en fonction de leur probabilité de détection et requiert donc des informations sur les caractéristiques liées à la probabilité d'un résultat positif. Ces informations comprennent les facteurs affectant l'exposition et des informations sur la fréquence de ces facteurs au sein de la population, comme la proportion de personnes dans chaque catégorie d'exposition (11).

La probabilité d'un résultat positif peut également inclure des facteurs inhérents aux données de la surveillance passive (12). Les deux principales stratégies de collecte d'échantillons de COVID-19 au Canada sont passives dans le sens où les personnes testées ont décidé de se rendre dans un centre de santé parce qu'elles ont développé des symptômes ou sont des personnes à risque, préoccupées par l'exposition. En revanche, une stratégie de dépistage volontaire est une surveillance active dans le sens où elle consiste à rechercher des personnes à tester. D'autres études traitent des stratégies permettant de prendre en compte le biais de sous-estimation lorsque toutes les personnes malades ne se présentent pas pour des soins de santé, dans le contexte de l'intégration de données de surveillance passive et active (12–14).

Au début d'une maladie émergente, il se peut que l'on ne dispose pas d'informations suffisantes pour tenir compte des difficultés d'utilisation des données issues de la surveillance passive. L'objectif de cette étude d'intervention est de présenter une stratégie de surveillance ciblée, qui combine des échantillons de surveillance passive et active, et qui utilise des informations minimales pour déterminer le nombre d'échantillons à recueillir par unité de surface et par unité de temps afin de détecter de nouveaux agrégats de cas de COVID-19 — à un stade suffisamment précoce pour permettre l'isolement des cas, la recherche des contacts et leur mise en quarantaine — afin de contrôler une éclosion.



Méthodes

Pour déterminer la nécessité d'un échantillonnage volontaire, la première étape consiste à déterminer si le nombre prévu d'échantillons obtenus à partir de la surveillance des infections virales respiratoires fondée sur les soins de santé fournit une taille d'échantillon déjà adéquate pour détecter de nouveaux agrégats de cas de COVID-19 au seuil de prévalence d'infection souhaité dans la population générale pour la période d'intérêt. Si la taille de l'échantillon s'avère insuffisante, la deuxième étape consiste à déterminer combien d'échantillons hebdomadaires supplémentaires sont nécessaires chez des personnes volontaires pour détecter de nouveaux agrégats de cas de COVID-19 au seuil de prévalence souhaité dans la population générale.

Nous avons utilisé une approche pondérée relative, dans laquelle le niveau de prévalence attendu dans une section particulière de la population définit la pondération que cet échantillon aurait dans la détection de la COVID-19 à p_0 . Cette approche suppose un échantillonnage aléatoire au sein du groupe d'échantillonnage. Chaque échantillon reçoit des points de pondération compte tenu de la prévalence attendue dans son groupe de population. La collecte des échantillons se poursuit jusqu'à ce que le nombre de points soit suffisant pour détecter la COVID-19 à p_0 . Nous avons démontré notre méthode à l'échelle P/T et à une fréquence hebdomadaire, bien que cette approche puisse être adaptée à d'autres unités régionales ou périodes de temps.

Étape 1 : Déterminer si suffisamment d'échantillons sont prélevés sur les patients symptomatiques dans les établissements de soins de santé

Avant la COVID-19 au Canada, le dépistage des virus respiratoires était destiné aux patients hospitalisés, ainsi qu'aux établissements et aux épidémies, où il aurait le plus d'impact sur les soins cliniques (15). Cependant, le dépistage de la COVID-19 est désormais recommandé pour toutes les personnes symptomatiques au Canada (16). Ici, les données sur les visites médicales avant la COVID-19 pour les personnes présentant des symptômes d'infections respiratoires sont utilisées pour déterminer le nombre attendu de visites médicales hebdomadaires au cours desquelles le dépistage de la COVID-19 pourrait avoir lieu.

Pour les données relatives aux visites médicales effectuées avant la pandémie de COVID-19, nous devons choisir une période récente au cours de laquelle aucune autre pandémie n'était en cours. Pendant la pandémie de grippe H1N1 en 2009–2010, il y a eu manifestement plus de consultations médicales pour des symptômes d'infection virale que la plupart des autres années. Nous avons supposé que si la COVID-19 est contrôlée à un niveau de risque acceptable, le nombre attendu de visites sera

conforme aux visites médicales des années autres que celles où la pandémie de grippe H1N1 a eu lieu. Par conséquent, nous avons utilisé le nombre annuel moyen de visites signalées pour la période non pandémique de 2016 à 2018 comme moyenne des visites médicales annuelles prévues pour le Canada ($n = 13\,310\,000$) (tableau 1; Institut canadien d'information sur la santé, analyse non publiée pour l'Agence de la santé publique du Canada, 2020). Le nombre attendu de consultations médicales hebdomadaires par P/T, E , peut alors être calculé en fonction de la taille de la population de la P/T et de l'unité de temps :

Équation 1 :

$$E = \frac{\text{taille de la population P/T}}{\text{taille de la population canadienne}} \times \frac{\text{nombre annuel de visites au Canada}}{52 \text{ semaines}}$$

Tableau 1 : Estimation du nombre annuel de visites de soins ambulatoires et d'admissions pour maladie respiratoire pendant une période non pandémique, Canada, 2016–2018

Type de visite	Nombre de visites
Admissions à l'hôpital	220 000 ^a
Visites aux services d'urgence	1 900 000 ^a
Visites aux soins de santé primaires	11 000 000 ^a
Nombre de résidents dans les foyers de soins de longue durée ^b	190 000
TOTAL	13 310 000

^a Institut canadien d'information sur la santé (ICIS). Nombre annuel, moyenne des exercices fiscaux 2016–2018

^b Institut canadien d'information sur la santé, 2020; fait référence aux foyers de soins de longue durée financés/subventionnés par l'État

Pour déterminer si E est suffisant pour détecter la COVID-19 le plus tôt possible à un niveau de risque acceptable, il est nécessaire de définir le niveau de prévalence seuil, p_0 , dans la population générale pour détecter et contrôler l'éruption de nouveaux cas. À titre de référence, l'Allemagne a utilisé un niveau de $p_0 = 0,05\%$ au cours de son processus de levée des mesures de confinement (1). Ce niveau correspond à une prévalence sur sept jours de 50/100 000. Nous étudions ici une valeur plus prudente de $p_0 = 0,025\%$ correspondant à une prévalence sur sept jours de 25/100 000.

On s'attend à ce que les visites de personnes présentant des symptômes d'infections respiratoires aient une probabilité d'infection plus élevée que celles des personnes asymptomatiques. Nous supposons qu'une prévalence de 0,64 % dans la population des visites médicales est une valeur réaliste qui peut se produire lorsque la COVID-19 est contrôlée de manière acceptable et qu'il y a un assouplissement des mesures de santé publique. Cette valeur se situe dans la fourchette inférieure du pourcentage moyen de positivité hebdomadaire signalé dans le système d'analyse des tests de laboratoire (SALT) du Réseau canadien de renseignements sur la santé publique



(RCRSP) pour le mois de mai 2020, ce qui complète la période printanière où les mesures de santé publique maximales étaient en place au Canada. Ensuite, la pondération de la contribution des échantillons du groupe d'échantillonnage i , ici la population des visites médicales, avec une prévalence de p , pour détecter la COVID-19 à p_0 pendant la période d'intérêt t , est le suivant :

Équation 2 :

$$w(i, t) = p(i, t)/p_0$$

Cette pondération est ensuite utilisée pour traduire le nombre hebdomadaire de visites médicales E en nombre de points de pondération qui servent à détecter la COVID-19 :

Équation 3 :

$$wp(i, t) = \frac{E}{w(i, t)}$$

Le résultat, $wp(\text{soins de santé}, t)$, est ensuite comparé au nombre d'échantillons nécessaires pour détecter au moins un cas positif de COVID-19, $d(i, t)$, dans la population des visites médicales en utilisant un calcul normalisé de la taille de l'échantillon (4) :

Équation 4 :

$$d(i, t) = \frac{-\ln(1-\alpha)}{p \times f}$$

pour un $\alpha = 0,95$ étant la confiance de détecter au moins un cas positif de COVID-19 à un seuil de détection minimal $p = p(\text{soins de santé}, t)$, et $f = 0,79$ étant la sensibilité du test pour les échantillons provenant de personnes symptomatiques (17). La taille de l'échantillon augmentera avec des niveaux croissants de α . Les valeurs typiques se situent entre 0,95 et 0,99, et à mesure que l'on dispose de plus d'informations, il peut devenir évident que des niveaux plus élevés sont nécessaires pour détecter la transmission communautaire suffisamment tôt pour contrôler l'écllosion. Si $wp(i, t) < d(\text{soins de santé}, t)$, alors davantage d'échantillons provenant de membres du public ne se rendant pas dans un centre de soins de santé sont nécessaires pour détecter au moins un cas positif de COVID-19 à p_0 .

Étape 2 : Déterminer combien d'échantillons hebdomadaires supplémentaires sont nécessaires de la part des volontaires

S'il n'y a pas assez d'échantillons de visites médicales, une deuxième étape est utilisée pour calculer combien d'échantillons supplémentaires sont nécessaires dans la population générale pour une détection précoce pendant la période d'intérêt t . L'équation 4 est utilisée à nouveau, mais cette fois-ci du point de vue de l'utilisation d'un échantillonnage de volontaires pour détecter la COVID-19 à p_0 , ce qui signifie que dans l'équation 4, $p = p_0$. De plus, les volontaires sont pour la plupart asymptomatiques, nous définissons donc une sensibilité de test plus faible $f = 0,70$ pour les personnes asymptomatiques (18,19). Le résultat, $d(\text{volontaire}, t)$, est ensuite utilisé pour calculer le nombre de tests supplémentaires nécessaires auprès des

volontaires, compte tenu de l'effort d'échantillonnage parmi les visites médicales, E , comme suit :

Équation 5 :

$$a(t) = d(\text{volontaire}, t) - E$$

Pour optimiser la collecte d'échantillons auprès des volontaires, nous appliquons l'approche de pondération relative pour cibler l'échantillonnage en fonction de la probabilité de dépistage positif. La sélection des groupes de volontaires dépend de la connaissance et de la disponibilité des données relatives aux caractéristiques qui influencent la probabilité d'un positif à la COVID-19. Dans la démonstration de notre méthode, nous avons défini des groupes de volontaires par niveau de taux de contact selon les données relatives à la profession (*données non publiées du Centre d'information sur le marché du travail, Statistique Canada, à la demande de l'Agence de santé publique du Canada. 2020*), bien que d'autres caractéristiques de données puissent également être utilisées (e.g. l'historique des déplacements, le groupe d'âge). L'hypothèse est que le fait de cibler l'échantillonnage sur des groupes à plus haut risque réduit la taille globale de l'échantillon nécessaire pour détecter la COVID-19. Nous créons ici trois populations de volontaires plausibles dont la prévalence d'infection attendue diffère selon le nombre de contacts (faible, moyen et élevé) qu'ils ont avec d'autres personnes (collègues de travail ou autres membres du public) chaque jour en fonction de leur profession. Pour cet exemple, nous utilisons une prévalence pour le contact moyen de 0,04 %. Il s'agit de la prévalence moyenne observée en Alberta chez les personnes asymptomatiques qui n'étaient pas des contacts étroits ou ne faisaient pas partie des enquêtes sur les éclosions pendant une période allant du 14 février au 5 juillet 2020 (*données non publiées du gouvernement de l'Alberta, 2020*). Nous supposons que les niveaux de prévalence des groupes de contact faible et élevé sont alors respectivement le double et la moitié de ceux du groupe de contact moyen.

La prévalence du groupe d'échantillonnage i est utilisée pour calculer la pondération de sa contribution, $w(i, t)$, à la détection de la COVID-19 à p_0 en utilisant l'équation 2. Ensuite, le nombre de tests nécessaires dans chaque population de volontaires, en plus de E , pour détecter au moins un cas positif de COVID-19 à p_0 étant donné $w(i, t)$ est calculé comme suit :

Équation 6 :

$$v(i, t) = \frac{\alpha(t)}{w(i, t)}$$

La valeur de $v(i, t)$ est le nombre total d'échantillons à tester si l'on échantillonne exclusivement dans ce groupe. La dernière considération consiste à calculer le nombre optimal d'échantillons de tests nécessaires pour tous les groupes d'échantillonnage volontaires, compte tenu de la probabilité d'échantillonnage de leur population. Les données de l'Enquête sur la population active de mars 2020 (20) et de la base de données sur les professions O*Net (*données non publiées*



du Centre d'information sur le marché du travail, Statistique Canada, à la demande de l'Agence de la santé publique du Canada. 2020) définissent la proportion de Canadiens ayant un emploi avec un taux de contact faible, moyen et élevé, *proportion(i)*, comme étant respectivement de 0,112, 0,392 et 0,494. Ainsi, la probabilité qu'un échantillon de test provienne du groupe d'échantillonnage volontaire *i*, dans une province ou un territoire à *t*, étant donné qu'il ne fait pas partie de *E* est :

Équation 7 :

$$Pr(i,t) = \lambda \times proportion(i)$$

où λ est la probabilité de ne pas faire partie de la population des visites médicales : $1 - E/P/T$ *taille de la population*. Par conséquent, le nombre total d'échantillons de tests nécessaires provenant de toutes les populations volontaires dans une province ou un territoire à *t* pour détecter au moins un cas positif de COVID-19 à p_0 est :

Équation 8 :

$$Z(t) = \sum_i^J v(i, t) \times Pr(i, t)$$

Où *i* est le groupe d'échantillonnage volontaire et *J* est le nombre total de groupes d'échantillonnage.

Cette méthode dépend de la taille de la population compte tenu du calcul de *E*. Pour évaluer la sensibilité de la taille de la population, nous montrons également les résultats pour *Z(t)* lorsque $p_0 = 0,05 \%$, afin de comparer la proportion de la population qui doit faire l'objet d'une surveillance lorsque $p_0 = 0,05 \%$ et $p_0 = 0,025 \%$.

Résultats

Nous présentons ici les résultats pour la taille de l'échantillon déterminée à l'échelle provinciale et à une fréquence hebdomadaire. Pour tous les provinces et territoires, nous avons supposé les mêmes niveaux de prévalence pour les groupes d'échantillonnage. Si l'on considère uniquement la pondération de la contribution à la détection de la COVID-19 à p_0 compte tenu de la prévalence supposée des groupes d'échantillonnage, les échantillons provenant de la population des visites médicales ont au moins huit fois plus de chances de donner un résultat positif au test de COVID-19 (i.e. 25,6/3,20) (**tableau 2**).

Comme cela est inhérent au calcul, les provinces et territoires dont la population est plus importante auront un nombre plus élevé de consultations médicales attendues, *E*. Étant donné la pondération élevée de la contribution de cette population pour détecter la COVID-19 à p_0 , les populations plus importantes nécessiteront moins d'échantillons supplémentaires, voire aucun, pour la détection précoce. Si l'objectif est de détecter la COVID-19 à p_0 à l'échelle des provinces et territoires pendant

Tableau 2 : Niveaux de prévalence et pondération des groupes d'échantillonnage de volontaires par rapport à la population des visites médicales avec des taux de contact faibles, *l*, moyens, *m*, et élevés, *h*

Groupes échantillons, <i>i</i>	Prévalence, $p(i, t)$	Poids, $w(i, t)$
Visites aux services des soins de santé	0,64	25,6
Volontaires ayant un taux de contact élevé	0,08	3,20
Volontaires ayant un taux de contact moyen	0,04	1,60
p_0	0,025	1,0
Volontaires ayant un taux de contact faible	0,02	0,80

la période d'intérêt *t*, alors seules les provinces de la Colombie-Britannique, l'Alberta, l'Ontario et le Québec auraient un nombre suffisant d'échantillons de visites médicales (**tableau 3**). Cela suppose des visites pour des maladies respiratoires aux niveaux de prévalence supposés lorsque les mesures de santé publique maximales étaient en place, de la mi-mars jusqu'à juste avant la période de leur assouplissement en mai 2020.

Tableau 3 : Identification des provinces et territoires qui manquent d'échantillons pour la population de visites médicales^{a,b}

Province/territoire	<i>E</i> ^c	<i>wp(soins de santé, t)</i>	<i>d(soins de santé, t)</i>
CB	34 522	1 349	593
AB	29 809	1 164	
SK ^d	7 982	312	
MB ^d	9 304	363	
ON	99 371	3 882	
QC	57 668	2 253	
NB ^d	5 268	206	
NS ^d	6 602	258	
PE ^d	1 068	42	
NL ^d	3 522	138	
YK ^d	277	11	
NT ^d	303	12	
NV ^d	264	10	

Abréviations : AB, Alberta; CB, Colombie-Britannique; MB, Manitoba; NB, Nouveau-Brunswick; NL, Terre-Neuve-et-Labrador; NS, Nouvelle-Écosse; NT, Territoires du Nord-Ouest; NV, Nunavut; PE, Île-du-Prince-Édouard; ON, Ontario; QC, Québec; SK, Saskatchewan; YK, Yukon

^a Identification des provinces et territoires qui manquent d'échantillons pour la population des visites médicales, sur la base du nombre de visites médicales prévues, *E*, converti en points de pondération, *wp(soins de santé, t)*, et comparé au nombre d'échantillons pondérés *d(soins de santé, t)* nécessaire pour détecter la COVID-19 dans la population des visites médicales à p_0 .

^b Les valeurs sont arrondies vers le haut

^c Le nombre attendu d'échantillons provenant des visites médicales à l'échelle P/T, *E*, et ce nombre traduit en points de pondération pour la détection de la COVID-19, *wp(soins de santé, t)*

^d Identification des provinces et territoires qui manquent d'échantillons pour la population de visites médicales



À l'étape 2, on constate que les groupes d'échantillonnage à faible taux de contact nécessitent des échantillons modèles pour une détection précoce (tableau 4). Dans le calcul du nombre optimal d'échantillons supplémentaires nécessaires pour détecter la COVID-19 à p_0 lorsqu'on ajoute des échantillons de volontaires, $Z(t)$, le faible nombre de E par rapport à la population totale de la province ou du territoire fait que $Pr(i,t)$ est très proche de la proportion de personnes ayant des professions avec des taux de contact faibles, moyens et élevés (tableau 5). Les provinces et territoires moins peuplés ont besoin d'un plus grand nombre d'échantillons volontaires pour la détection précoce, car leur E est plus faible, et donc le pourcentage de la population qui doit se porter volontaire est plus élevé. Lorsque p_0 passe de 0,25 % à 0,50 %, le Manitoba dispose d'un nombre suffisant d' E pour la détection précoce et le pourcentage de la population nécessitant un prélèvement volontaire dans les autres provinces et territoires est réduit de moitié (tableau 5).

Tableau 4 : Nombre d'échantillons nécessaires pour détecter la COVID-19^{a,b}

Province/ territoire	$d_{volontaire}(t)$	$n_{volontaire}(t)$	Taux de contact faible	Taux de contact moyen	Taux de contact élevé
SK	17 118	9 137	11 421	5 711	2 855
MB		7 814	9 767	4 884	2 442
NB		11 850	14 812	7 406	3 703
NS		10 516	13 145	6 573	3 286
PE		16 050	20 063	10 031	5 016
NL		13 597	16 996	8 498	4 249
YK		16 841	21 051	10 526	5 263
NT		16 815	21 019	10 509	5 255
NV		16 854	21 068	10 534	5 267

Abréviations : MB, Manitoba; NB, Nouveau-Brunswick; NL, Terre-Neuve-et-Labrador; NS, Nouvelle-Écosse; NT, Territoires du Nord-Ouest; NV, Nunavut; PE, Île-du-Prince-Édouard; SK, Saskatchewan; YK, Yukon

^a Nombre d'échantillons nécessaires afin de détecter la COVID-19 à p_0 pour la sensibilité du test aux personnes asymptomatiques, $d_{volontaire}(t)$; nombre de tests nécessaires en plus des échantillons de visites médicales provenant de tous les groupes d'échantillonnage de volontaires, $n_{volontaire}(t)$, et si l'on échantillonne exclusivement dans chaque groupe avec des contacts faibles, moyens et élevés au travail

^b Les valeurs sont arrondies vers le haut

Tableau 5 : Nombre optimal d'échantillons supplémentaires nécessaires pour détecter la COVID-19

P/T	i	Taille de la population	E	λ	Proportion (i)	$Pr(i,t)$	$wp(i,t)$	$Z(t)$	% P/T	$Z(t)$ à $p_0 = 0,05\%$	P/T en % à $p_0 = 0,05\%$
SK	Faible	1 181 666	7 982	0,99	0,112	0,111	11 420	4 867	0,41	307	0,26
	Moyen			0,99	0,392	0,386	5 710				
	Élevé			0,99	0,494	0,490	2 855				
MB	Faible	1 377 517	9 304	0,99	0,112	0,111	9 766	4 162	0,30	s.o.	s.o.
	Moyen			0,99	0,392	0,386	4 883				
	Élevé			0,99	0,494	0,490	2 441				
NB	Faible	779 993	5 268	0,99	0,112	0,111	14 812	6 312	0,81	1 753	0,23
	Moyen			0,99	0,392	0,386	7 406				
	Élevé			0,99	0,494	0,490	3 703				
NS	Faible	977 457	6 602	0,99	0,112	0,111	13 144	5 602	0,57	1 042	0,11
	Moyen			0,99	0,392	0,386	6 572				
	Élevé			0,99	0,494	0,490	3 286				
PE	Faible	158 158	1 068	0,99	0,112	0,111	20 063	8 550	5,41	3 991	2,52
	Moyen			0,99	0,392	0,386	10 031				
	Élevé			0,99	0,494	0,490	5 016				
NL	Faible	515 828	3 522	0,99	0,112	0,111	17 042	7 263	1,41	2 703	0,52
	Moyen			0,99	0,392	0,386	8 521				
	Élevé			0,99	0,494	0,490	4 261				
YK	Faible	41 078	277	0,99	0,112	0,111	21 051	8 972	21,8	4 412	10,7
	Moyen			0,99	0,392	0,386	10 526				
	Élevé			0,99	0,494	0,490	5 263				
NT	Faible	44 904	303	0,99	0,112	0,111	21 019	8 958	20 0	4 398	9 80
	Moyen			0,99	0,392	0,386	10 509				
	Élevé			0,99	0,494	0,490	5 255				



Tableau 5 : Nombre optimal d'échantillons supplémentaires nécessaires pour détecter la COVID-19 (suite)

P/T	<i>i</i>	Taille de la population	<i>E</i>	λ	Proportion (<i>i</i>)	$Pr(i, t)$	$wp(i, t)$	$Z(t)$	% P/T	$Z(t)$ à $p_0 = 0,05\%$	P/T en % à $p_0 = 0,05\%$
NV	Faible	39 097	264	0,99	0,112	0,111	21 068	8 979	23 0	4 419	11 3
	Moyen			0,99	0,392	0,386	10 534				
	Élevé			0,99	0,494	0,490	5 267				

Abréviations : MB, Manitoba; NB, Nouveau-Brunswick; NL, Terre-Neuve-et-Labrador; NS, Nouvelle-Écosse; NT, Territoires du Nord-Ouest; NV, Nunavut; PE, Île-du-Prince-Édouard; P/T, province/territoire; SK, Saskatchewan; s.o., sans objet; YK, Yukon

Remarque : À p_0 , lors de l'augmentation avec des échantillons de volontaires du groupe d'échantillonnage i , $Z(t)$, et les valeurs sous-jacentes pour le calcul, y compris λ , la probabilité de ne pas faire partie de la population attendue de visites médicales, E . Le pourcentage de la population provinciale qui devrait participer au dépistage volontaire à l'unité de temps t est également indiqué

Discussion

Nous présentons une approche pondérée relative pour calculer le nombre d'échantillons de tests nécessaires pour détecter au moins un cas de COVID-19 à un niveau seuil pour la détection précoce et le contrôle de nouvelles éclosions. Cette approche combine les nombres de tests attendus de visites médicales, avec un échantillonnage supplémentaire dans la population générale. Parmi les groupes d'échantillonnage, la probabilité de détecter la COVID-19 est la plus élevée dans la population de patients se rendant à une visite médicale. Lorsque les échantillons de ce groupe sont insuffisants, l'échantillonnage de la population générale à l'aide d'une approche de pondération relative peut fournir les échantillons supplémentaires requis.

Notre approche est plus réalisable pour les grandes populations, car elles présentent des taux de dépistage plus élevés de la population de visites médicales. Si des échantillons supplémentaires sont nécessaires, la proportion de la population requise comme volontaires est plus réalisable qu'avec des populations plus petites. Par exemple, dans notre démonstration de la détermination de la taille de l'échantillon en utilisant les provinces et territoires comme population de surveillance, nous constatons que la Colombie-Britannique, l'Alberta, l'Ontario et le Québec ont déjà des tailles d'échantillon suffisantes dans la population des visites médicales à une fréquence hebdomadaire. L'augmentation des échantillons provenant de volontaires nécessite le dépistage de 0,3 à 0,81 % de la population pour la Saskatchewan, le Manitoba, le Nouveau-Brunswick et la Nouvelle-Écosse. Cependant, pour l'Île-du-Prince-Édouard, Terre-Neuve-et-Labrador, le Yukon, les Territoires du Nord-Ouest et le Nunavut, plus de 1 % à 23 % de la population doit être testée. Il est peu probable que le niveau de conformité ou la capacité à se rendre sur les sites de test soient atteints. Cette fourchette se réduit à 0,5 %–11,3 % de la population si l'on suppose que $p_0 = 0,05$ % au lieu de $p_0 = 0,025$ %, comme cela a été fait pour l'Allemagne après sa première vague d'infections par la COVID-19.

Forces et faiblesses

Nous n'avons pas pris en compte la spécificité du test dans notre approche. Les tests de réaction en chaîne par polymérase (PCR) pour l'infection par le SRAS-CoV-2 présentent une excellente

spécificité d'au moins 98 %, mais leur sensibilité est plus variable (21,22). Même à 98 %, les échantillons de grande taille peuvent donner lieu à un nombre considérable de faux positifs; par exemple, on peut s'attendre à 160 faux positifs en testant 8 000 personnes. Les résultats faussement positifs peuvent avoir des conséquences importantes si la personne ayant un résultat faussement positif subit un traitement qui serait non requis contre la COVID, pourrait mettre en danger sa santé. En revanche, un résultat faussement positif pour une personne en bonne santé n'entraînera qu'un isolement pendant un certain temps, ce qui aura un impact limité sur la santé de cette personne.

L'intérêt de notre approche est de guider les efforts de surveillance au début d'une maladie émergente lorsque l'on connaît peu les facteurs affectant la probabilité qu'un échantillon soit positif pour la maladie. Au début de l'émergence de la maladie, les systèmes de surveillance développent leur capacité à tester et à recueillir des informations pertinentes pour comprendre le risque de transmission. La collecte d'informations sur les facteurs à haut risque, tels que l'historique des voyages, accuse un retard par rapport aux informations sociodémographiques telles que le sexe et le groupe d'âge. En outre, l'association des informations sociodémographiques avec le résultat du test n'a peut-être pas encore été déterminée. Lorsque des informations sur les facteurs de risque élevés deviennent disponibles, les approches qui exploitent ce type d'informations dans une approche de pondération relative peuvent affiner les estimations de la détermination de la taille de l'échantillon, comme l'ont montré Jennelle et al. (10). Cette approche comprend la prise en compte des changements dans le risque de transmission au fil du temps, en fonction de l'augmentation, des pics et des diminutions du risque de la maladie. Il s'agit également de tenir compte de la nature passive des systèmes de surveillance, qui entraîne la violation de l'hypothèse selon laquelle l'échantillonnage est non aléatoire. Par exemple, les obstacles à l'accès aux soins de santé ou aux centres de dépistage en fonction du sexe, de l'âge, de la profession ou de l'origine ethnique. Par conséquent, la surreprésentation des personnes ayant un certain profil sociodémographique peut fausser l'exactitude des valeurs de prévalence pour les groupes d'échantillonnage. À l'heure actuelle, les échantillons prélevés par écouvillonnage



du nasopharynx chez les patients se rendant dans un centre de soins de santé primaires sont rarement effectués. Des méthodes d'échantillonnage moins invasives, telles que les tests de rinçage buccal, faciliteraient l'atteinte des tailles d'échantillon cibles.

Lors de l'émergence d'une nouvelle maladie, il est probable que les informations soient insuffisantes pour définir avec précision la probabilité d'un résultat positif au test, qui peut alors être utilisé pour déterminer la taille de l'échantillon pour une détection précoce. Nous présentons ici une méthode permettant d'estimer la taille des échantillons pour la détection précoce en utilisant des informations limitées, comme nous le montrons avec les niveaux de prévalence (à la fois estimés et supposés) de plusieurs groupes d'échantillonnage. Le fait de pondérer la contribution d'un échantillon provenant d'un groupe d'échantillonnage donné à l'obtention d'un résultat positif permet d'élaborer une stratégie d'échantillonnage plus efficace pour la détection précoce, ce qui aide à cibler les efforts et les ressources de surveillance.

Idéalement, les niveaux de prévalence sont actualisés, lorsque cela est possible, afin de réduire l'erreur dans les estimations de la taille de l'échantillon, car les niveaux de prévalence dans les groupes d'échantillonnage changent dans le temps et dans l'espace. Plus précisément, les P/T peuvent couvrir de vastes zones, où les villes peuvent être séparées par des centaines de kilomètres et, par conséquent, n'être que faiblement connectées en termes de facteurs d'infection. Il peut y avoir plusieurs unités épidémiologiques au sein d'une province ou territoire, ce qui signifie que les schémas de transmission communautaire sont plus similaires au sein d'une unité que parmi les unités voisines, et donc que les niveaux de prévalence dans les groupes d'échantillonnage peuvent différer entre les unités dans l'espace et dans le temps. Les mesures résultant de la surveillance, telles que la détermination de la taille de l'échantillon, sont idéalement réalisées au niveau spatial de l'unité épidémiologique (12). La méthode présentée ici peut être adaptée au niveau d'une unité épidémiologique. Cette approche garantirait que la détermination de la taille de l'échantillon pour la détection précoce reflète les efforts d'échantillonnage (i.e. E) et les niveaux de prévalence pour les groupes d'échantillonnage qui sont uniques à l'unité pendant la période d'intérêt.

Conclusion

Cette étude d'intervention présente une stratégie de surveillance ciblée, combinant des échantillons de surveillance passive et active, afin de déterminer le nombre d'échantillons à recueillir par unité de surface et par unité de temps pour détecter de nouveaux agrégats de cas de COVID-19. L'objectif de cette stratégie est de permettre une détection suffisamment précoce pour contrôler une éclosion.

Déclaration des auteurs

E. E. R. — Conception, analyse formelle, rédaction du projet original, rédaction-révision et édition
 R. R. — Conception, révision de l'écriture, revue critique
 N. H. O. — Conception, révision de l'écriture, revue critique

Intérêts concurrents

Aucun.

Remerciements

Aucun.

Financement

Aucun.

Références

1. Federal Government of Germany. A balanced outcome: Angela Merkel reports on meeting with state premiers. BREG; (modifié 2020-05-06; accédé 2021-03-24). <https://www.bundesregierung.de/breg-en/search/merkel-bund-laender-gespraech-1751090>
2. Ogden NH, Fazil A, Arino J, Berthiaume P, Fisman DN, Greer AL, Ludwig A, Ng V, Tuite AR, Turgeon P, Waddell LA, Wu J. Scénarios de modélisation de l'épidémie de COVID-19 au Canada. Relevé des maladies transmissibles au Canada 2020;46(6):225–31. DOI
3. Ng V, Fazil A, Waddell LA, Bancej C, Turgeon P, Otten A, Atchessi N, Ogden NH. Projected effects of nonpharmaceutical public health interventions to prevent resurgence of SARS-CoV-2 transmission in Canada. CMAJ 2020;192(37):E1053–64. DOI PubMed
4. Fosgate GT. Practical sample size calculations for surveillance and diagnostic investigations. J Vet Diagn Invest 2009;21(1):3–14. DOI PubMed
5. Hicks AL, Kissler SM, Mortimer TD, Ma KC, Tairaoa G, Ashcroft M, Williamson DA, Lipsitch M, Grad YH. Targeted surveillance strategies for efficient detection of novel antibiotic resistance variants. eLife 2020;9:e56367. DOI PubMed
6. Public Health Ontario. Enhanced epidemiological summary: COVID-19 in Ontario: A summary of wave 1 transmission patterns and case identification. PHO; 2020. <https://www.publichealthontario.ca/-/media/documents/ncov/epi/2020/08/covid-19-wave-1-transmission-patterns-epi-summary.pdf?la=en>
7. Short DL. Alberta Health no longer recommending asymptomatic testing. Edmonton Journal; (modifié 2020-09-17). <https://edmontonjournal.com/news/local-news/alberta-health-no-longer-recommending-asymptomatic-testing>
8. BC Centre for Disease Control. Phases of COVID-19 testing in BC. BCCDC; 2020. <http://www.bccdc.ca/health-info/diseases-conditions/covid-19/testing/phases-of-covid-19-testing-in-bc>



9. Walsh DP, Miller MW. A weighted surveillance approach for detecting chronic wasting disease foci. *J Wildl Dis* 2010;46(1):118–35. DOI PubMed
10. Jennelle CS, Walsh DP, Samuel MD, Osnas EE, Rolley R, Langenberg J, Powers JG, Monello RJ, memarest ED, Guber R, Heisey DM. Applying a Bayesian weighted surveillance approach to detect chronic wasting disease in white-tailed deer. *J Appl Ecol* 2018;55(6):2944–53. DOI
11. Dohoo I, Martin W, Stryhn H. *Veterinary Epidemiology Research* 2nd Edition. Charlottetown, Canada: VER Inc.; 2009.
12. Cameron AR, Meyer A, Faverjon C, Mackenzie C. Quantification of the sensitivity of early detection surveillance. *Transbound Emerg Dis* 2020;67(6):2532–43. DOI PubMed
13. Hadorn DC, Stärk KD. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Vet Res* 2008;39(6):57. DOI PubMed
14. Li X, Chang HH, Cheng Q, Collender PA, Li T, He J, Waller LA, Lopman BA, Remais JV. A spatial hierarchical model for integrating and bias-correcting data from passive and active disease surveillance systems. *Spat Spatio-Temporal Epidemiol* 2020;35:100341. DOI PubMed
15. Public Health Ontario. LABSTRACT – December 2020: Respiratory Virus Testing Update. PHO; 2020. <https://www.publichealthontario.ca/-/media/documents/lab/lab-sd-121-respiratory-viral-testing-algorithm-enhanced-surveillance-update.pdf?la=en>
16. Gouvernement du Canada. Document d'orientation national sur les indications de dépistage de la COVID-19 par tests réaction en chaîne de la polymérase (PCR). 2020 (accédé 2020-10). <https://www.canada.ca/fr/sante-publique/services/maladies/2019-nouveau-coronavirus/document-orientation/national-indications-depistage-laboratoire.html>
17. Kojima N, Turner F, Slepnev V, Bacelar A, Deming L, Kodeboyia S, Slausner JD. Self-Collected Oral Fluid and Nasal Swab Specimens Demonstrate Comparable Sensitivity to Clinician-Collected Nasopharyngeal Swab Specimens for the Detection of SARS-CoV-2. *Clin Infect Dis*. 2020;ciaa 1589. DOI
18. Public Health Ontario. COVID-19 Laboratory Testing Q&As. PHO; 2020. <https://www.publichealthontario.ca/-/media/documents/lab/covid-19-lab-testing-faq.pdf?la=en>
19. BC Centre for Disease Control. Interpreting the results of Nucleic Acid Amplification testing (NAT; or PCR tests) for COVID-19 in the Respiratory Tract. BCCDC; (modifié 2020-04-03). <https://medicalstaff.islandhealth.ca/sites/default/files/covid-19/testing/covid-interpreting-test-results-nat-pcr-bccdc.pdf>
20. Statistique Canada. Enquête sur la population active, mars 2020. Statistique Canada ; (modifié 2020-04-09). <https://www150.statcan.gc.ca/n1/en/daily-quotidien/200409/dq200409a-fra.pdf?st=kEd0dF4B>
21. Mustafa Hellou M, Górská A, Mazzaferri F, Cremonini E, Gentilotti E, De Nardo P, Poran I, Leeflang MM, Tacconelli E, Paul M. Nucleic acid amplification tests on respiratory samples for the diagnosis of coronavirus infections: a systematic review and meta-analysis. *Clin Microbiol Infect* 2021;27(3):341–51. DOI PubMed
22. Mironas A, Jarrom D, Campbell E, Washington J, Ettinger S, Wilbacher I, Endel G, Vrazic H, Myles S, Prettyjohns M. How to best test suspected cases of COVID-19: an analysis of the diagnostic performance of RT-PCR and alternative molecular methods for the detection of SARS-CoV-2. medRxiv. 2021.01.15.21249863. DOI