



Veille bibliographique de la COVID-19 – cadre pour gérer la littérature et appuyer la prise de décisions fondées sur des données probantes sur un sujet de santé publique en évolution rapide – Matériel supplémentaire

Tricia Corrin, Dima Ayache, Austyn Baumeister, Kaitlin Young, Kusala Pussegoda, Rukshanda Ahmad, Lisa Waddell

Méthodes

Leçons apprises par l'équipe d'analyse quotidienne et les ressources humaines

Afin d'assurer la continuité des analyses, trois responsables des analyses ont été formés pour s'acquitter des tâches de gestion de l'équipe d'examen, d'examen des articles, de tenue à jour de la base de données, de formation des nouveaux examinateurs et de compilation et de distribution du rapport d'analyse quotidien.

Pendant la durée de ce projet, il y a eu environ 50 examinateurs bénévoles de l'Agence de la santé publique du Canada (ASPC). Certains examinateurs se sont portés volontaires pendant quelques mois, tandis que d'autres ont travaillé à l'analyse quotidienne pendant la majeure partie du projet. Il a fallu environ 15 examinateurs travaillant quotidiennement depuis le début de la pandémie, soit environ six mois, pour traiter jusqu'à 1 000 articles par jour. En moyenne, un examinateur pouvait traiter un maximum de 50 citations primaires de recherche par jour. Étant donné que plusieurs des examinateurs ont donné de leur temps malgré d'autres priorités de travail, l'objectif de traitement était plutôt de 20 citations par jour. La capacité de l'Agence de la santé publique du Canada d'attirer des volontaires témoigne de la valeur de l'analyse pour plusieurs groupes dans l'ensemble du projet de l'Agence et de l'engagement de cette dernière à utiliser les données probantes dans la prise de décisions relativement à son travail d'intervention pour la pandémie. Sans l'adhésion et le soutien de l'ensemble de l'organisation pour ce projet, il n'aurait duré que quelques mois, car le volume de documentation aurait dépassé les ressources de la petite équipe de synthèse des connaissances.

Il était difficile d'obtenir le nombre requis d'examineurs pour effectuer les analyses quotidiennes. Plusieurs examinateurs formés et disponibles devaient contribuer au projet en tout temps. Il fallait donc gérer quotidiennement les examinateurs,

la planification et la formation fréquente des nouveaux examinateurs par les responsables des analyses.

Stratégie de recherche

Une recherche documentaire (publiée et prépubliée) a été effectuée quotidiennement du lundi au vendredi. Les recherches visant à extraire de la documentation pertinente sur la COVID-19 ont été effectuées dans sept bases de données : PubMed, Scopus, BioRxiv, MedRxiv, ArXiv, SSRN et Research Square. Un algorithme de recherche a été créé et ajusté en fonction des nouveaux termes utilisés dans la documentation. Cet algorithme a été adapté pour chaque base de données : COVID-19 OU SRAS-CoV-2 OU SRAS-coronavirus-2 OU nCov OU « nouveau CoV » OU (nouveau ET coronavirus). Lorsqu'elles étaient disponibles sur les serveurs de prépublication, on a également utilisé les collections liées à la COVID-19 pour trouver les articles pertinents.

La stratégie de recherche comprenait plusieurs serveurs de prépublication. Un seul, BioRxiv, avait une fonction de recherche et d'exportation pour les résultats de BioRxiv et de MedRxiv, semblable aux bases de données bibliographiques comme PubMed. Pour les autres, notamment ArXiv, SSRN et Research Square, les chercheurs devaient utiliser un processus plus manuel de recherche et d'extraction de citations. Par conséquent, ces dernières bases de données n'ont pas été incluses dans la plupart des autres bases de données documentaires sur la COVID-19 qui ont émergé pendant la pandémie.

Les documents prépubliés ont été utilisés plus que jamais pendant la pandémie, car il s'agissait des sources de données émergentes les plus récentes sur la situation pandémique, le virus et la maladie. En concevant notre base de données, de nombreuses vérifications croisées ont été menées avec des pages de renvoi d'éditeurs comme Lancet, Elsevier, The New England Journal of Medicine, BMJ, Wiley, Springer Nature et ChinaCDCweekly. De plus, la base de données a été comparée à d'autres analyses documentaires (e.g. le projet de documentation mondiale de l'OMS) et la saisie de



la base de données a fait l'objet de vérifications continues pendant les examens rapides sur des sujets précis en vérifiant si nos chercheurs avaient omis une recherche en particulier. À l'occasion, des citations omises ont été relevées, notamment parce que la citation n'était indexée dans aucune de nos bases de données ou qu'une citation avait été accidentellement supprimée pendant le processus de déduplication.

Gestion de base de données

Afin de faciliter l'accès à l'ensemble de la base de données, les utilisateurs finaux ont reçu une liste quotidienne de sujets avec des points saillants, un lien vers les références dans [REFWORKS](#) et un lien vers les rapports quotidiens compilés en [format Excel interrogeable](#).

Classement de la documentation sur la COVID-19

Au début de ce projet, deux outils de classement par intelligence artificielle (DistillerAI) ont été formés pour classer automatiquement la documentation sur la COVID-19 en trois groupes définis : 1) la recherche de synthèse (e.g. examens systématiques, méta-analyses, examens de la portée); 2) la recherche primaire; et 3) la recherche secondaire (e.g. examens narratifs, commentaires, lettres aux rédacteurs). Chaque jour,

après avoir utilisé les outils de classement DistillerAI pour trouver les résultats de recherche, les responsables de l'analyse ont effectué la sélection de tout article que DistillerAI n'a pas été en mesure de classer. Des recherches secondaires ont été compilées et vérifiées pour s'assurer d'un minimum d'erreurs de classification avant d'être ajoutées au rapport final.

Les examinateurs devaient remplir un formulaire préalable qui comprenait des renseignements pour chaque article de recherche primaire. En moyenne, un examinateur rédigeait de 10 à 50 formulaires d'articles par jour. Les membres de l'équipe ont rédigé un résumé pour chaque article afin de communiquer le but de l'étude et les principaux résultats. Les résumés étaient nécessaires, car plusieurs articles n'en comprenaient pas, ou comprenaient des résumés trop longs. Un ensemble de thèmes a été élaboré et défini pour les examinateurs afin que les articles puissent être classés de façon uniforme et étiquetés par sujets d'intérêt. Au total, il y avait 21 thèmes parmi lesquels les examinateurs pouvaient choisir des articles. Il s'agissait notamment de l'épidémiologie, de la transmission, des interventions en santé publique, de la coronavirologie, de la recherche sur les vaccins et des thérapies. Le **tableau S1** fournit une liste complète des thèmes et de leurs définitions.

Tableau S1 : Thèmes – Terminologie et définitions

Thèmes	Définitions
Modélisation/prévision	La modélisation prédictive est un processus qui utilise l'exploration des données et la probabilité pour prévoir les résultats. Chaque modèle est composé de plusieurs prédicteurs, qui sont des variables susceptibles d'influencer les résultats futurs.
Épidémiologie	Traite de l'incidence, de la distribution et du contrôle possible des maladies et d'autres facteurs liés à la santé. Il s'agit notamment de la variable R_0 , du taux d'attaque, du temps de doublement du nombre de cas, du taux de mortalité, de l'intervalle de série, du taux d'attaque clinique, de la fraction asymptomatique, et de la proportion des cas asymptomatiques et infectieux.
Transmission	Passage d'un pathogène causant une maladie transmissible d'un hôte infecté à un hôte non infecté par des voies directes ou indirectes.
Données cliniques	Données sur les cas : Il s'agit notamment des paramètres cliniques comme la période d'incubation, la période latente, la période de communicabilité, la durée de la maladie, la durée de l'hospitalisation, les facteurs de risque de l'hôte, ainsi que les profils cliniques des patients (les symptômes présentés au cours de la maladie, les comorbidités).
Séquelles à long terme	Données ou études qui décrivent les effets à long terme de la COVID-19. Il s'agit notamment de symptômes persistants (ou chroniques) d'essoufflement, de fatigue, de maux de tête, de douleurs musculaires ou articulaires, de lésions aux organes (e.g. poumons, cœur, reins), de caillots sanguins, de vaisseaux sanguins et d'autres problèmes décrits comme étant survenus avec la maladie ou persistants après la fin de la maladie aiguë de la COVID-19. Acronymes : COVID longue, aussi appelée COVID chronique, ou le syndrome post-COVID-19.
Surveillance	Collecte, analyse et interprétation continues et systématiques de données sur la santé. Les données de surveillance seront probablement le nombre de cas soupçonnés et confirmés, le nombre de décès et le nombre de cas rétablis.
Coronavirologie	Recherche relative au virus : ses caractéristiques, sa composition génétique, les analyses phylogénétiques.
Diagnostic et détection des agents pathogènes	Études sur l'identification du virus; culture, tests PCR, tests d'anticorps et antigéniques, etc.
Thérapeutique	Études de substances pouvant être utilisées pour traiter les hôtes infectés, notamment les produits d'immunisation passive.
Recherche sur les vaccins	Études de candidats vaccins pour prévenir l'infection par le SRAS-CoV-2, notamment des essais cliniques.
Économie	Toutes les études liées à l'économie et au coût d'une écloison.
Modèle animal	Toutes les études utilisant des modèles animaux; habituellement un sous-domaine d'un autre domaine.
Immunologie	Études mesurant la réponse immunitaire; souvent un sous-domaine de données cliniques.
Interventions en santé publique	Toutes les études évaluant l'efficacité d'une intervention de santé publique donnée (dans le cas d'un modèle prédictif); e.g. distanciation physique, contrôle à l'aéroport.



Tableau S1 : Thèmes – Terminologie et définitions (suite)

Thèmes	Définitions
Santé mentale	Toutes les études sur les répercussions ou les conséquences de la pandémie sur la santé mentale qui mettent l'accent sur la santé mentale pour la population en général et les travailleurs de la santé.
Intervention de la santé publique	Les articles dans cette catégorie comprennent la COVID-19 ET la toxicomanie, la violence familiale, l'analyse des médias sociaux, le respect des mesures de santé publique et les études sur les connaissances, les attitudes et les croyances liées à la pandémie.
Priorités en matière de santé publique	Ces citations mettent l'accent sur les priorités de recherche actuelles ou sur les lacunes dans les connaissances.
PCI/IPAC	Prévention et contrôle des infections/Infection Prevention and Control. La recherche sur l'efficacité des interventions de PCI devrait également être étiquetée comme PCI.
Réponse de la santé	Toutes les interventions et procédures liées aux soins de santé ou les connaissances, l'attitude et les comportements des travailleurs de la santé. D'autres articles portent sur l'efficacité du triage des patients en fonction du niveau de soins, la description des changements majeurs aux soins de santé (e.g. changements dans les admissions aux soins d'urgence ou changements dans les chirurgies et répercussions sur les chirurgies ou les traitements retardés en raison de la pandémie), etc.
ECR	Essais cliniques randomisés et leurs protocoles.
Zoonose	Tous les documents traitant de la transmission ou de l'occurrence d'une infection naturelle au SRAS-CoV-2 chez les animaux.
AUTRE	Plusieurs sujets d'étude ne s'appliquaient pas aux autres thèmes. Cette catégorie comprend des articles sur la sécurité alimentaire, les répercussions sur les entreprises, les changements dans les habitudes de vie (e.g. l'alimentation, l'activité physique, la sexualité, les habitudes d'observation des oiseaux, etc.), les changements dans le nombre d'accidents de voiture, les répercussions sur les pertes d'emploi, les répercussions environnementales du confinement, et des articles sur la politique et la pandémie.
EXCLUSIONS	Il s'agit des articles qui ne portaient pas sur la COVID-19 et ses conséquences directes, par exemple, la recherche sur les styles d'éducation et les programmes d'études, les rapports de cas qui ne sont pas liés à la COVID-19, les documents du marché boursier, les chirurgies à l'ère de la COVID-19 (e.g. rapports de cas ou une seule intervention chirurgicale qui a eu lieu pendant la pandémie, mais qui n'est pas liée à un cas de COVID-19) OU de toute autre intervention qui n'était pas liée à la pandémie et à la santé publique.
Examen de la documentation	Les articles qui résument la documentation publiée devraient être étiquetés comme un examen. Il s'agit notamment des examens systématiques, des méta-analyses, des examens de la portée, des aperçus et des examens généraux (la liste n'est pas exhaustive).
Commentaires/éditoriaux	Les commentaires, les éditoriaux, les lettres à la rédaction et les autres types d'articles d'opinion pour lesquels il n'y a pas de nouvelles données recueillies par l'auteur ou d'analyse effectuée par l'auteur doivent être inscrits dans la catégorie des commentaires. Toutes les « lignes directrices et recommandations » et tout article qui ne contient pas de données primaires devraient être classés dans la catégorie des examens ou des commentaires.
Nouvelles	Il s'agit des études sans données primaires et sans pertinence scientifique.

Abréviations : COVID-19, maladie à coronavirus 2019; ECR, essai clinique randomisé; IPAC, Infection Prevention and Control; PCI, Prévention et contrôle des infections; PCR, réaction en chaîne par polymérase; SRAS-CoV-2, coronavirus 2 du syndrome respiratoire aigu sévère; TS, travailleur de la santé