Study Protocol

Long Title: Language translation of knowledge mobilization resources: A randomized trial

Short Title: Language translation of knowledge mobilization resources

Trial registration: NCT ID not yet assigned

Protocol Version 1

Funding

This project is funded in part by an Alberta Innovates Summer Studentship

Roles and Responsibilities

Principal Investigators: Dr. Lisa Hartling and Dr. Sarah Elliott

Project Coordinators: Samantha Cyrkot, and Chelseay Robles

Lay Summary

Knowledge mobilization (KM) resources are tools designed to facilitate the use of research evidence in healthcare decision-making. These resources are created in various formats – including plain language summaries, infographics, and videos – to meet the needs of diverse end-users, such as healthcare professionals, policymakers, patients, and caregivers. They are intended to be easily accessible; however, individuals whose first language is not English may have difficulty understanding them. Thus, translating KM resources into other languages is essential to support health equity and accessibility, but it is often costly and time intensive.

This study aims to explore whether artificial intelligence (AI) tools, specifically ChatGPT - an AI-based large language model developed by OpenAI - can effectively translate KM resources for members of the public whose first language is not English. The resource being evaluated offers guidance on preventing post-COVID-19 condition and has already been translated by a professional (human) translator into seven languages commonly spoken in Canada: French, Spanish, Ukrainian, Tagalog, Arabic, Chinese, and Punjabi. Using ChatGPT, AI-generated translations will be created in those same seven languages.

For this study, participants – adults living in Canada whose first language is one of the selected languages and able to read English – will be randomly assigned to review either an AI-generated or a professionally translated version of a KM resource. They will then complete a questionnaire evaluating their understanding of the resource, as well as the readability and acceptability of the translation.

This study will contribute to our understanding of the potential use of AI for translating health information. Our goal is to support equitable access to health information and promote citizen-centered care by reducing language barriers using innovative solutions.

Background and Rationale

KM resources play a critical role in bridging the gap between research and practice in healthcare. Effective communication of this information is essential to ensure the uptake and implementation of research evidence and health recommendations by the public.

Canada is home to a highly diverse population. According to the 2021 Canadian Census, immigrants make up 23% of the population and this figure is projected to increase to 29–34% by 2041.[2,3] In line with this trend, many Canadian households speak a language other than English or French at home. Over 13% of the national population primarily use a non-official language at home, and 21% of private households are multilingual (at least two languages are spoken within their households).[4] In Alberta specifically, Tagalog (4%), Punjabi (3%), and Mandarin (2%) are among the most frequently spoken non-official languages.[1] These three languages represent some of the key communities of interest for this project.

Language proficiency remains a persistent challenge for many newcomers. According to the Longitudinal Survey of Immigrants to Canada (LSIC), 26% of new immigrants continued to report difficulties with English or French four years after arrival.[5] Inadequate health communication stemming from language barriers can lead to poorer health outcomes, reduced satisfaction with care, and increased healthcare inequalities.[6] High-quality translations of health materials are therefore essential to promoting equitable access to health information and helping reduce disparities in health outcomes. While professional translation services have long been the standard for producing linguistically accurate translated materials, they are often resource-intensive, limiting the scalability and timeliness of current translation efforts.

Recent advances in AI have generated interest in leveraging AI-powered tools to support the translation of health materials. AI has the potential for fast, scalable translation that could reduce costs and improve timely access to multilingual resources. Among the many AI translation tools, ChatGPT has emerged as a large language model with demonstrated strengths in semantic accuracy, fluency, and adaptability across multiple languages.[7,8] Studies suggest that ChatGPT can produce translations that are comparable to human translators in many contexts, suggesting it generates natural, contextually appropriate language.[9] Additionally, ChatGPT is publicly accessible, requiring only standard web access, lowering barriers for healthcare providers, organizations, and individuals seeking rapid, on-demand translation support. Its broad adoption has also resulted in a growing body of research on its capabilities, limitations, and applications in healthcare communication and beyond.[7,10-12] These factors informed the decision to focus on ChatGPT for this study.

Nevertheless, concerns remain regarding the ability of AI powered tools to navigate the cultural, contextual, and emotional nuances of language that are vital in healthcare communication. Translation extends beyond the literal meaning of words, its cultural nuances, contextual relevance, and the social and emotional connotations embedded in language.[12] AI-generated translations may misrepresent meanings, overlook cultural sensitivities, or fail to convey the appropriate tone leading to controversial interpretations, consequences that could compromise trust and comprehension among end-users.[13,14] To our knowledge, no controlled trials have yet

compared AI and professional translations specifically for KM resources aimed at the public. Empirical assessments of AI-generated translations are therefore essential for understanding its capabilities and limitations.

By investigating professional versus AI-generated translation strategies, this study will provide critical insights into the potential role of AI in enhancing the accessibility and effectiveness of health communication for linguistically diverse populations. Ultimately, the findings will inform future strategies for using innovative technologies to support equitable, multilingual access to healthcare information and contribute to the development of more inclusive health communication strategies in Canada.

Objectives

The overarching research question is: Can AI effectively translate KM resources for the public? For this study, the primary objective is to compare KM resources translated using an AI tool versus those translated by professional human translators. Specifically, we hypothesize that translations of KM resources produced by ChatGPT will be comparable to professional human translations in terms of understanding, readability, and acceptability. This study aims to inform and provide evidence for the implications of AI for communication of health information.

Methods

Patient and Public Involvement

Members of the Pediatric Parent Advisory Group (P-PAG) will provide input throughout the study.[15] They will give input on study design, assist with pilot testing data collection instruments, advise on recruitment strategies, and promote the study within their networks. They will assist with interpreting results and will critically contribute to developing and disseminating findings (e.g., plain language summary of results).

Trial Design

This study is a randomized controlled trial. Participants will be randomly assigned (1:1 allocation ratio) to receive either an AI-translated version or a professionally translated version of a KM resource. They will be asked to answer a set of questions about their understanding of the content of the resource, its readability and acceptability.

Trial Setting

The trial will be conducted entirely online. Participants will complete the study remotely using their own personal devices (e.g., computers or tablets) with internet access. All materials, including study link, instructions, and questionnaires, will be delivered electronically.

Eligibility Criteria

Inclusion Criteria

- 18 years of age or older
- Living in Canada

- First language is one of the selected languages (i.e. French, Spanish, Ukrainian, Tagalog, Arabic, Chinese, and Punjabi)
- Able to read and complete a questionnaire in English
- Access to an electronic device (e.g. computer or tablet), Internet and email

Exclusion Criteria

- Under the age of 18 years
- Does not live in Canada
- First language is not one of the selected languages (i.e. French, Spanish, Ukrainian, Tagalog, Arabic, Chinese, and Punjabi)
- Unable to read and complete a questionnaire in English
- No access to an electronic device (e.g. computer or tablet), Internet or email

Intervention and Comparator

Participants will be randomized to view a KM resource translated, either by AI (i.e., ChatGPT) or a professional (human) translator, into one of the selected languages. The KM resource (an online resource sheet) was developed by our group for the Public Health Agency of Canada about the prevention of post-covid condition (PCC). The resource presents recommendations that were developed by an international group according to standards for guideline development. The resource has already been professionally translated into French as well as six languages common to newcomers across Canada (Arabic, Punjabi, Ukrainian, Tagalog, Chinese, Spanish). The English-version of the KM resource is included in Appendix A and is available online (https://canpcc.ca/app/uploads/2025/03/Online-Resource-Sheet-GT1-Prevention-of-PCC-Final.pdf).

Outcomes

The following outcomes will be used to evaluate the translated versions of the KM resource from multiple dimensions: knowledge acquisition (understanding), ease of text processing (readability), and confidence in the information (acceptability).

The primary outcome will be:

Understanding: Measured by participant responses to a set of 7 multiple-choice questions specifically designed to evaluate key content knowledge about post COVID-19 condition and related recommendations contained in the resource. The number of correct answers will serve as an objective indicator of understanding, assessing the ability of each translation method to preserve critical concepts.

Secondary outcomes will be:

Readability: assessing readability is critical because even an accurate word-for-word translation may fail to be effective if the participants find the text difficult to read or process. Evaluated using a 5-point Likert scale assessing clarity, grammar, ease of reading, and terminology, alongside additional open-ended questions identifying confusing or unfamiliar vocabulary.

Acceptability: evaluated through participants' trust in the translation. This includes 5-point Likert-scale ratings of cultural and linguistic appropriateness, comfort in using the resource, and trust in the information presented. Participants will also indicate perceptions about the translation method (human or AI). Acceptability reveals participant confidence and willingness to use the KM resource, which are key determinants of successful health messaging and adherence to recommendations.

A draft of the online survey questions is provided in Appendix B.

Harms

This study involves little to no risk to participants. Nevertheless, potential harms may include mild fatigue or frustration if participants find the text confusing or difficult to interpret. To minimize this, all materials will be pilot-tested for clarity and presented in a user-friendly format. Participants may withdraw from the study at any time without penalty.

Participant timeline

Participants in this study will be involved in a single-session conducted online. All data will be collected immediately after the participant has reviewed the assigned translation. Each participant will complete the study in approximately 10-15 minutes. No long-term follow-up is planned.

Sample Size

Due to the pilot nature of this study, we aim to recruit 50 participants per language group, with 25 participants randomized to evaluate the AI-translated version and 25 assigned to evaluate the professionally translated version. With seven language groups, the total target sample size will be 350 participants. This sample size follows guidance for exploratory and pilot studies that require no formal sample size calculation. With 50 participants per language group, we will be able to provide estimates of potential differences while allowing us to explore variations by language.

Recruitment

Recruitment will begin once the study has received ethics approval. Potential participants will be recruited through a variety of online strategies to ensure broad and diverse enrollment across language groups. Recruitment efforts will include targeted electronic newsletters distributed via University of Alberta's undergraduate and graduate student digests, as well as outreach to international student groups. Additionally, a recruitment email will be sent to our pre-established connections to interest-holder groups across Canada (e.g. OurKidsHealth, TREKK, Cochrane Canada) to inform them about the study and to ask them to share the study materials within their networks. This may also include asking them to share via traditional means of communication (e.g. e-newsletter, list serv).

Potential participants (self-identified) will be able to access the study directly from the recruitment material by visiting the study link provided. The study recruitment materials will

also include email contact information (arche@ualberta.ca). If they email expressing interest in the study, the project coordinator will respond and provide a link to the study. The study link will include detailed instructions to complete the following:

- Answer eligibility screening questions
- Identify their first language
- Review Study Information letter
- Read KM resource to which they have been randomized (AI or professional translation in participant's first language)
- Complete the questionnaire including demographics and responses to questions regarding the KM resource

Incentives: Participants will be compensated $5 CAD. This incentive will come in the form of an electronic CAD gift card to the participant via Everything Gift Card.

Recruitment will occur through publicly shared recruitment materials (e.g., email digests, electronic flyers, etc.) containing a study description and a link to the study. Interested individuals will self-identify as potentially eligible and access the study directly by clicking the link. Upon accessing the study link, participants will be presented with a brief set of screening questions to confirm eligibility. Those who do not meet the inclusion criteria will be automatically thanked for their interest and exited from the survey. At the end of the survey, personal information (first name and email) will be shared (if the participant chooses to do so) with the researchers on a separate form, unlinked to the survey responses, for the purposes of gift card distribution. Information collected will be housed separately from the study data in the Gift Card Log.

Randomization and Blinding

Participants will be stratified by their first language then randomly assigned (1:1 allocation ratio) to one of two groups: the AI-translated KM resource or the professionally translated version. Randomization will be performed at the individual participant level using a computer-generated random allocation sequence to ensure unbiased group assignment.

Participants will be blinded to the type of translation they receive. They will be informed that the study compares different translation methods but will not be told whether the materials they view were translated by AI or by professional translators.

Data Collection and Management

Data will be collected and managed using REDCap (Research Electronic Data Capture). REDCap has an extensive privacy policy (https://help.redcap.ualberta.ca/policy-procedure/privacy) that has been reviewed by the study team and determined suitable for this study, and has been previously approved for research use at the University of Alberta. Data will be downloaded regularly and stored long-term on a secure server in the PI's faculty (Faculty of Medicine & Dentistry) at the University of Alberta. Upon completion of the study, the data will be archived on the Faculty of Medicine & Dentistry server. The Faculty of Medicine & Dentistry

has secure network drives accessible by network linked, password-protected computers and the data will not be accessible to anyone outside of the research team. Identifiable electronic data (e.g., Gift card records) will be deleted at time of study completion. De-identified (anonymous) electronic data (questionnaire data) will be kept for 5 years. Participants' email addresses and first names will be collected for purposes of sending the electronic gift card. No hard copies of the data files will be stored.

Statistical Methods

We will collect demographic data (e.g., age, gender, ethnicity, education level, country of birth, health literacy, and proficiency in languages spoken) to describe our study population. We will use descriptive statistics (numbers, frequencies, means with standard deviations) to analyze and present demographic data. We will determine overall understanding by calculating a score (ranging from 0 to 7) for each participant based on the number of questions answered correctly. We will calculate a mean score for each group and compare groups (professional vs. AI translation) using a two-sample t-test assuming equal variances. Other outcomes will be compared between groups using independent t-test or chi squared tests dependent on the type of data (i.e., continuous or categorical). Statistical uncertainties will be expressed with 95% confidence intervals; $p<0.05$ will indicate statistical significance. Analyses will be conducted based on intention-to-treat, i.e., all available data will be included and we will not know whether or the extent to which participants read the translated version of the KM tool.

Data monitoring and committee

Not applicable to this study.

Trial monitoring

No on-site monitoring is planned, as all data collection and oversight will occur remotely. The study coordinator and research team will review REDCap data weekly and data will be downloaded regularly (e.g., weekly) for quality control, cleaning and analysis. This data will be stored securely on an encrypted network drive at the University of Alberta, accessed via a password protected computer. Measures like adding a captcha will be used to prevent fraudulent entries. Recruitment and completion rates will also be monitored to inform outreach/recruitment strategies.

Ethics

Research Ethics Approval

Prior to the start of this trial, ethics approval will be sought from the University of Alberta's Health Research Ethics Board (REB). The study protocol, recruitment documents, and all participant materials will be submitted for review to ensure compliance with ethical standards and regulatory requirements. No participant recruitment or data collection will begin until formal approval has been obtained.

Protocol Amendments

The Principal Investigator (PI) will have authority regarding protocol amendments. Changes, including those affecting participant safety, study conduct, or data integrity, will be promptly communicated to the REB, trial registries, and all relevant interest-holders. Updated protocol versions will be documented accordingly.

Informed Consent

Participants will receive comprehensive information regarding the study's purpose, procedures, risks, benefits, confidentiality measures, and their rights, including the right to withdraw at any time without penalty. Consent will be obtained through implied consent by overt action. We will notify each participant that by clicking the link to complete the screening questions, reviewing the Study Information Letter, and completing the questionnaire, will mean that they provide implied consent by overt action to participate in the study. Participants will be able to download a copy of the full Study Information Letter for their own records.

Confidentiality

This study is designed to maintain participant anonymity and confidentiality at all times. Participants will not be asked to provide any identifying information as part of the survey itself, therefore, survey responses will remain anonymous. All data will be collected through a secure online platform and stored on password-protected systems accessible only to the research team. Although complete anonymity is intended, participants will be informed that confidentiality cannot be absolutely guaranteed due to the nature of internet-based data collection (e.g., theoretical risks related to platform security). However, no identifying information will be collected unless a participant voluntarily provides it (e.g., to receive a gift card), in which case this information will be stored separately from survey responses.

Ancillary and Post-Trial Care

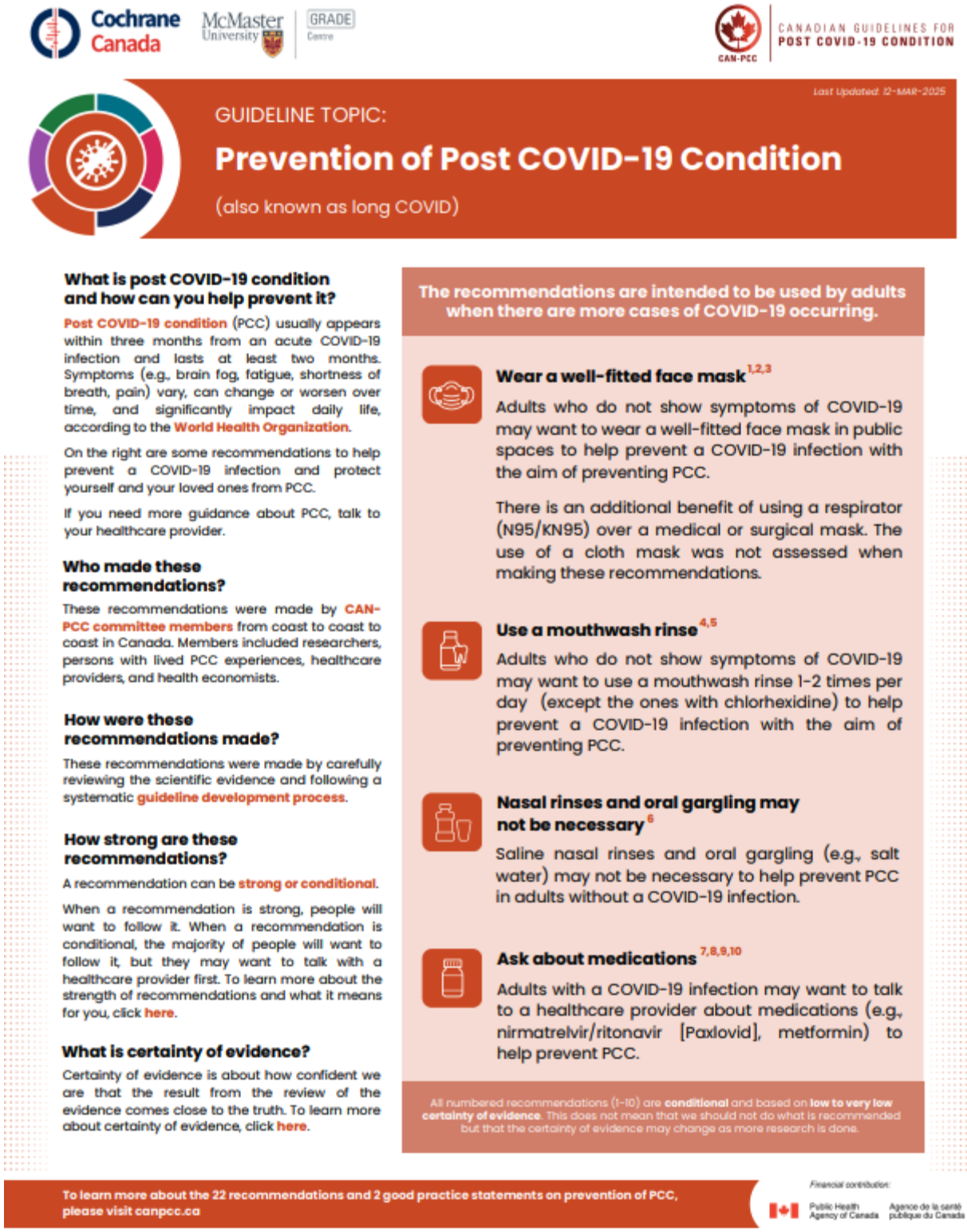
Although this trial involves minimal risk to participants, any concerns or adverse effects identified during the study will be addressed promptly. Participants will be provided with contact information for the study team to report any issues.

References

1. Statistics Canada. Census of Population. Immigrant population by non-official language spoken most often at home, Alberta. 221.
2. Statistics Canada. Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as Canadians [Internet]. Statistics Canada. Government of Canada; 2022 [cited 2025 Jun]. Available from: https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm
3. Government of Canada. The Daily - Canada in 2041: A larger, more diverse population with greater differences between regions [Internet]. www150.statcan.gc.ca. 2022 [cited 2025 Jun]. Available from: https://www150.statcan.gc.ca/n1/daily-quotidien/220908/dq220908a-eng.htm
4. Government of Canada SC. Multilingualism of Canadian households [Internet]. www12.statcan.gc.ca. 2023 [cited 2025 Jun]. Available from: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021014/98-200-X2021014-eng.cfm
5. Schellenberg G, Maheux H. Canadian Social Trends, Special Edition 2007. Ottawa, ON: Statistics Canada; 2007. [Accessed 2007 October 4]. Immigrants' perspectives on their first four years in Canada: highlights from three waves of the longitudinal survey of immigrants to Canada. Available from: http://dsp-psd.pwgsc.gc.ca/collection_2007/statcan/11-008-X/11-008-XIE20070009627.pdf
6. Whitaker KL, Krystallidou D, Williams ED, Black G, Vindrola-Padros C, Braun S, et al. Addressing language as a barrier to healthcare access and quality. British Journal of General Practice [Internet]. 2022 Jan;72(714):4–5. Available from: https://bjgp.org/content/72/714/4
7. Genovese A, Sahar B, Gomez-Cabello CA, Haider SA, Srinivasagam P, Forte AJ, et al. Artificial intelligence in clinical settings: a systematic review of its role in language translation and interpretation. Annals of Translational Medicine [Internet]. 2024 Dec 1;12(6):117–7. Available from: https://atm.amegroups.org/article/view/132363/html
8. Grimm DR, Lee YJ, Hu K, Liu L, Garcia O, Balakrishnan K, et al. The utility of ChatGPT as a generative medical translator. European Archives of Oto-Rhino-Laryngology. 2024 May 5; Available from: https://link.springer.com/content/pdf/10.1007/s00405-024-08708-8.pdf
9. Rao P, McGee LM, Seideman CA. A Comparative assessment of ChatGPT vs. Google Translate for the translation of patient instructions. Journal of Medical Artificial Intelligence. 2024 Jun;7:11–1. Available from: https://cdn.amegroups.cn/journals/jlpm/files/journals/33/articles/9019/public/9019-PB1-7668-R1.pdf
10. Meddeb A, Lüken S, Busch F, Adams L, Ugga L, Koltsakis E, et al. Large Language Model Ability to Translate CT and MRI Free-Text Radiology Reports Into Multiple Languages. Radiology. 2024 Dec 1;313(3). https://doi.org/10.1148/radiol.241736

11. Garcia OA, Charat Thongprayoon, Jadlowiec CC, Mao SA, Napat Leeaphorn, Pooja Budhiraja, et al. Advancing health equity: evaluating AI translations of kidney donor information for Spanish speakers. Frontiers in Public Health. 2025 Jan 27;13. https://doi.org/10.3389/fpubh.2025.1484790
12. Moneus AM, Sahari Y. Artificial intelligence and human translation: A contrastive study based on legal texts. Heliyon [Internet]. 2024 Mar 30;10(6):e28106. Available from: https://www.sciencedirect.com/science/article/pii/S2405844024041379#abs0010
13. Kreienbrinck A, Hanft-Robert S, Forray AI, Nozewu A, Mösko M. Usability of technological tools to overcome language barriers in healthcare– a scoping review. Archives of Public Health. 2025 Feb 25;83(1). https://doi.org/10.1136/bmjopen-2023-079814
14. Moy S, Irannejad M, Manning SJ, Farahani M, Ahmed Y, Gao E, et al. Patient Perspectives on the Use of Artificial Intelligence in Health Care: A Scoping Review. Journal of Patient-Centered Research and Reviews [Internet]. 2024 Apr 2;11(1):51–62. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11000703/
15. Hartling L, Elliott SA, Buckreus K, Leung J, Scott SD. Development and evaluation of a parent advisory group to inform a research program for knowledge translation in child health. Research Involvement and Engagement. 2021 Jun 7;7(1). https://doi.org/10.1186/s40900-021-00280-3

Appendix A – KM online resource sheet



Appendix B – Summary of outcomes and draft questions for online surveys

Outcomes	**Scoring**
Understanding (i.e., comprehension of the information contained in the infographic)	Percentage of correct answers (0-7)
1. What is post COVID-19 condition (PCC)? a) A condition that resolves within two weeks of infection **b) A condition with ongoing symptoms for at least two months after COVID-19 infection** c) A condition caused by an allergic reaction to COVID-19 vaccines d) A condition only affecting individuals who were hospitalized due to complications of COVID-19 e) I'm not sure 2. Which of the following is recommended to prevent a COVID-19 infection and reduce the risk of PCC? a) Avoiding crowded outdoor areas b) Taking vitamin supplements daily **c) Wearing a well-fitted face mask** d) Taking over-the-counter medications after exposure e) I'm not sure 3. What does a *conditional recommendation* mean? a) A recommendation only supported by strong evidence and should be followed by most people when making health decisions b) A recommendation that only applies during specific situations, such as public health lockdowns or restrictions c) A recommendation that only applies to certain provinces in Canada, depending on local healthcare policies and needs **d) A recommendation that people may follow, but they might want to consult a healthcare provider before making decisions** e) I'm not sure 4. The resource mentions that all numbered recommendations are based on low to very low certainty of evidence. What does this mean? a) The recommendations are not evidence-based and should not be followed at this time b) The information has not been fully reviewed or confirmed by experts	

c) The recommendations are useful, but they may change based on future research d) The recommendations are not meant to be followed e) I'm not sure 5. Based on the resource, which of the following is true about the group that developed the recommendations? **a) It included people with lived experience, healthcare providers, researchers, and economists** b) It was developed based on responses from public polling and surveys c) It was led entirely by academic researchers, public health officials, and policymakers outside Canada d) It included only public health officials from provincial governments e) I'm not sure 6. According to the resource, why might someone with no COVID-19 symptoms still consider using a mouthwash rinse (excluding chlorhexidine)? a) To reduce bacteria in the mouth that may cause dental problems over time b) To follow a recommendation issued by public health authorities across country **c) To reduce the chance of becoming infected with COVID-19 and possibly prevent PCC** d) To relieve throat discomfort that may be caused by COVID-19 infection e) I'm not sure 7. What was NOT considered in the guideline when recommending mask types? a) The difference between respirators and surgical masks b) The benefit of N95/KN95 respirators c) The role of public mask-wearing in preventing PCC **d) The use of cloth masks** e) I'm not sure	

Readability (i.e., is it easy to read and process)	5-point Likert scale: 5 'strongly agree', 3 'neither agree or disagree', 1 'strongly disagree' Open comment boxes
Likert Scale Statements 1) The resource was easy to read. 2) The language used in the resource was easy to understand. 3) The sentences and grammar used in the resource made sense to me. 4) I had no trouble following the main ideas in the resource. 5) The text used in the resource flowed well. Please elaborate on your answers *(text box)* Open-ended questions 1. Is there any strange or unfamiliar vocabulary? Please list any words or phrases in the resource that were difficult to understand. *(text box)* 2. Was any part of the translation confusing or unclear? Please write down any suggestions to improve the clarity or wording of the translation? *(text box)*	
Method of Translation and Acceptability	5-point Likert scale: 5 'strongly agree', 3 'neither agree or disagree', 1 'strongly disagree' Multiple choice – single answer Open comment boxes
Likert Scale Statements 1) The translation was appropriate for my cultural and linguistic background. 2) I understood the recommendations provided in the resource. 3) I feel comfortable using this resource for health information. 4) I trust the information provided in this resource. Please elaborate on your answers *(text box)* Single Answer 1. Do you think this resource was translated by: a. A professional human translator? b. An Artificial Intelligence (AI) tool (such as ChatGPT or Google Translate)?	

2. Please explain your answer in question 5. *(text box)* 3. Does knowing whether the translation was done by a human or artificial intelligence (AI) affect your trust in the resource? a. Yes, I would trust a human translator more b. Yes, I would trust an AI translator more c. No, I do not trust AI d. No, the source does not affect my trust e. I'm not sure 4. Please explain your answer in question 7. *(text box)* 5. Was the translation acceptable to you as a version to share public health information? a. Yes b. No c. I'm not sure 6. Please explain your answer in question 9. *(text box)* 7. Would you prefer to read this resource in English instead of your first language? a. Yes b. No 8. Please explain your answer in question 11. *(text box)* 9. Is there anything else you would like to share with us about the resource or the translation? *(text box)*	