Canada's Drug and
Health Technology Agency

CADTH Methods and Guidelines

Guidance for Reporting Real-World Evidence

May 2023

ISSN: 2563-6596

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

Preamble

A key component of CADTH's 2022 to 2025 strategic plan is to be a leader in the practice of evidence appraisal, including real-world evidence (RWE).[1] CADTH has partnered with Health Canada, the Institut national d'excellence en santé et en services sociaux (INESSS), and other health system stakeholders to advance the integration of RWE into decision-making. RWE is defined as the evidence surrounding the usage and potential benefits or risks of a medical product, derived from analysis of real-world data (RWD).[2] The use of RWD as a primary or complementary data source to generate RWE is a possible way to reduce evidence uncertainties. RWE provides an opportunity to include consideration of health technologies and potential other benefits that lie outside those observed with traditional clinical trials, which can provide additional context to the recommendations that facilitate decision-making. Regardless of evidence type or data source, the principles outlined in this document highlight the importance of transparency in reporting to help ensure the credibility of the evidence.

Prospectively planned randomized controlled trials (RCTs) continue to be the most robust tool for providing evidence of drug safety and efficacy. However, while RCTs facilitate the collection of high-quality controlled data, generalizability in the real world is limited. RCTs often do not address all research questions relevant to assessments of comparative clinical effectiveness. Additionally, the conduct of RCTs is not always feasible or ethical for certain diseases or disorders (such as some rare diseases) or patient populations (such as children, pregnant people, or older adults).

CADTH and Health Canada are the co-chairs of the RWE Steering Committee, which includes pan-Canadian health, government, and patient organizations; industry; academia; and data holders. The RWE Steering Committee provided support and oversight for this initiative and received quarterly updates from the Guidance for Reporting Real-World Evidence working group (WG). The Guidance for Reporting Real-World Evidence WG included a Methods Authorship Team, a Leadership Review Team, and an Expert Methods Panel including Stakeholder Panel members.

Guidance for Reporting Real-World Evidence lays the foundation for the use of RWE in regulatory approval and Health Technology Assessment (HTA) in Canada, starting with the principles for reporting of RWE studies. CADTH, Health Canada, and INESSS intend to use the guidance as appropriate for their individual needs, aligned to the principles outlined in the document. This initiative forms the foundation for transparent reporting of RWE studies in Canada and facilitates appraisal of RWE for the purpose of supporting decision-making.

The overall purpose of this guidance is to promote standardization in the reporting of RWE studies for those undertaking and submitting RWE studies of health technologies to support decision-making in Canada. The specific objectives were to:

- identify existing global guidance on principles and standards for reporting on RWE studies to create an initial draft of Canadian RWE reporting standards that align with international standards
- establish consensus on items to be included in the core reporting standards for Canadian RWE studies through engagement with national and international experts in RWD and RWE.

A Methods Authorship Team embarked on a multistep, multistakeholder process to develop *Guidance for Reporting Real-World Evidence* in collaboration with the Leadership Review Team, which included representatives from CADTH, Health Canada, and INESSS. This process included mapping key concepts identified via a review of existing RWE guidance, conducting an adapted Delphi process with an Expert Methods Panel of Canadian and international experts, and facilitating discussions between experts including representatives of key Canadian health system organizations, including CADTH, Health Canada, INESSS, CIHI and Statistics Canada. A wide variety of stakeholders were engaged iteratively to provide feedback on the guidance, through a public consultation period that included a series of meetings, public information sessions, and an 8-week stakeholder feedback period. A *Response to Stakeholder Feedback* report outlines the feedback received during this consultation process by theme and the modifications that were made to the draft guidance.

This document will be most relevant to those developing and reporting RWE to regulatory and HTA bodies, as well as those who review and appraise evidence. *Guidance for Reporting Real-World Evidence* outlines principles that are intended to be consistent with regulatory and HTA standards both in Canada and internationally. The expectation is that this document will be periodically updated or expanded over time as the field of RWD and RWE evolves.

Guidance for Reporting Real-World Evidence Authors and Contributors

Guidance for Reporting Real-World Evidence was created through a collaborative effort of health system leaders and stakeholders including academics, methodologists, health care providers, HTA organizations, regulators, payors, and data holders.

Oversight and support of the development of *Guidance for Reporting Real-World Evidence* was organized by the RWE Steering Committee, co-chaired by CADTH (Nicole Mittmann) and Health Canada (Kelly Robinson).

The Guidance for Reporting Real-World Evidence WG was co-chaired by CADTH (Laurie Lambert), Health Canada (Andrew Raven), and an invited RWE Methods Expert (Mina Tadrous), who provided quarterly updates to the RWE Steering Committee. The Guidance for Reporting Real-World Evidence WG included a Methods Authorship Team, a Leadership Review Team, and an Expert Methods Panel including the Stakeholder Panel members. Mina Tadrous was the lead of the Guidance Methods Authorship Team, which conducted the literature review and developed the guidance principles presented in this document.

The content of this report does not necessarily always reflect the views of the organizations acknowledged in the sections that follow.

Real-World Evidence Steering Committee

Co-Chairs
Nicole Mittmann, MSc, PhD (CADTH)

Kelly Robinson, MSc (Health Canada)

CADTH
Farah Husein, BScPhm, PharmD, MSc (HEOR)

Laurie Lambert, MPH, PhD

Tarry Ahuja, PhD

Abera Surendran, PhD

Health Canada
Melissa J. Hunt, MSc

Melissa Kampman, BSc, MSc, PhD

Michelle Mujoomdar, PhD

Catherine Njue, PhD

Andrew Raven, MSc

Craig Simon, PhD

Canadian Drug Agency Transition Office (CDATO)
Barry Jones, BPharm

Julie Robert, BSc

Institut national d'excellence en santé et en services sociaux (INESSS)
Geneviève Plamondon, MSc

Sylvie Bouchard, BPharm, DPH, MSc, MBA

pan-Canadian Pharmaceutical Alliance (pCPA)
Daniel Sperber, MSc, BHSc

Dominic Tan, BBA

Canadian Association of Provincial Cancer Agencies (CAPCA)
Gunita Mitera, PhD

Canadian Organization for Rare Disorders (CORD)
Durhane Wong-Rieger, PhD

Canadian Institute of Health Research (CIHR)
Jennifer Campbell, BSc, MA

Étienne Richer, PhD

Canadian Institute for Health Information (CIHI)
Deborah Cohen, PhD

Jordan Hunt, BComm, MA

Statistics Canada
Erik Dorff, BSc

Scott McLeish, MSc

Health Data Research Network Canada (HDRN Canada)
James Ted McDonald, PhD, MCom

Nicole Yada, MSc

Innovative Medicines Canada – BIOTECanada (IMC-BTC)
Innovative Medicines Canada and BIOTECanada had representation on the RWE Steering Committee who were engaged and consulted during the development of this guidance. CADTH acknowledges the input and guidance of all members of the RWE Steering Committee, including the representatives of Innovative Medicines Canada and BIOTECanada.

Invited Specialist Experts
Mina Tadrous, PharmD, MS, PhD

Kelvin Kar-Wing Chan, MD, MSc, PhD

Guidance for Reporting Real-World Evidence WG

Methods Authorship Team
Mina Tadrous, PharmD, MS, PhD, Leslie Dan Faculty of Pharmacy, University of Toronto (lead)

Christine Fahim, MSc, PhD, St. Michael's Hospital, Unity Health Toronto

Kaley Hayes, PharmD, PhD, Brown University School of Public Health

Theresa Aves, MSc, Institute of Health Policy, Management and Evaluation, University of Toronto

Leadership Review Team

CADTH

Nicole Mittmann, MSc, PhD

Farah Husein, BScPhm, PharmD, MSc (HEOR)

Laurie Lambert, MPH, PhD

Health Canada

Kelly Robinson, MSc

Melissa Kampman, BSc, MSc, PhD

Craig Simon, PhD

Institut national d'excellence en santé et en services sociaux (INESSS)

Sylvie Bouchard, BPharm, DPH, MSc, MBA

Geneviève Plamondon, MSc

Expert Methods Panel

Canadian Members

Erin C. Strumpf, PhD, McGill University, Québec City, Quebec

James Ted McDonald, PhD, MCom, University of New Brunswick, Fredericton, New Brunswick

Jason Robert Guertin, PhD, Université Laval, Québec City, Quebec

Jeff Round, BA (Hons), MA, PhD, University of Alberta, Edmonton, Alberta

Kelvin Kar-Wing Chan, MD, MSc, PhD, University of Toronto, Toronto, Ontario, and CanREValue Collaboration

Lisa Lix, BSHEc, MSc, PhD, PStat, University of Manitoba, Winnipeg, Manitoba

Mary A. De Vera, PhD, University of British Columbia, Vancouver, British Columbia

Robert Platt, PhD, McGill University, Montréal, Quebec

Sanja Stanojevic, PhD, Dalhousie University, Halifax, Nova Scotia

Scott Klarenbach, MD, MSc, University of Alberta, Edmonton, Alberta

International Members

Dalia Dawoud, BSc, MSc, PhD, National Institute for Health and Care Excellence (NICE), UK

Daniel Prieto-Alhambra, MD, PhD, University of Oxford, Oxford, England

Donna R. Rivera, PharmD, MSc, FDA, Maryland, US

Seamus Kent, PhD, National Institute for Health and Care Excellence (NICE) (*at the time of contribution*), UK

Shirley V. Wang, PhD, ScM, Brigham Women's Hospital, Harvard Medical School, Massachusetts, US

Canadian Stakeholder Members

Canadian stakeholder members participated in discussions at all Expert Methods Panel meetings, reviewed documents, and provided the perspective of RWE experts from key Canadian stakeholder organizations in the Canadian health system.

CADTH

Amanda Allard, PhD, Director, Pharmaceutical Reviews

Cody Black, MSc, Lead, Health Economics

Hongbo Yuan, PhD, Scientific Advisor, Scientific Advice, Methodologies, and Resources

Mike Innes, MSc, PharmD, Scientific Advisor, Scientific Advice, Methodologies, and Resources

Sheri Pohar, BScPharm, MScPharm, PhD, Scientific Advisor, Scientific Advice, Methodologies, and Resources

Health Canada

Andrew Raven, MSc, Manager, Biostatistics, Epidemiology, and Pharmacometrics Unit, Prescription Drugs Directorate

Catherine Njue, PhD, Manager, Office of Biostatistics, Biologic and Radiopharmaceutical Drugs Directorate

Institut national d'excellence en santé et en services sociaux (INESSS)

Geneviève Plamondon, MSc, Professionnelle scientifique, Bureau — Méthodologies et éthique

Sara Beha, MSc, Coordonnatrice clinique, Unité d'évaluation des technologies innovantes en santé, Direction de l'évaluation des médicaments et des technologies à des fins de remboursement

Naji-Tom Samaha, BSc, BPharm, MSc, Professionnel scientifique pharmacien, Direction de l'évaluation des médicaments et des technologies à des fins de remboursement

Canadian Institute for Health Information (CIHI)

Roger Cheng, BScPharm, PharmD, Program Consultant, Pharmaceuticals, Data Management Team

Statistics Canada

Erik Dorff, BSc, Chief, Chief Chronic Conditions and Symptoms, Centre for Population Health Data

Acknowledgements

The *Guidance for Reporting Real-World Evidence* team would like to acknowledge the important contribution of Barry Jones to this project. His insights and knowledge have helped advance the science and practice of reporting RWE for health technology assessments and regulatory approval.

Table of Contents

Abbreviations **14**
Summary **15**
What Is RWE? 15
Overall Purpose and Main Objectives 16
Specific Objectives 16
About This Guidance 16
Implementation Considerations 17
Background and Methods 17
Overview and Structure 18
Section 1: Research Questions and Study Design **19**
Overview 19
Specific Considerations and Recommendations 20
Section 1: Summary of Recommendations 21
Section 2: Setting and Context **22**
Overview 22
Specific Considerations and Recommendations 22
Section 2: Summary of Recommendations 23
Section 3: Data Specifications: Access, Cleaning Methods, and Linkage **24**
Overview 24
Specific Considerations and Recommendations 24
Section 3: Summary of Recommendations 25
Section 4: Data Sources, Data Dictionary, and Variables **26**
Overview 26
Specific Considerations and Recommendations 26
Section 4: Summary of Recommendations 27
Section 5: Participants **28**
Overview 28
Specific Considerations and Recommendations 29

Section 5: Summary of Recommendations 31

Section 6: Exposure Definitions and Comparators 31

Overview 31

Specific Considerations and Recommendations 32

Section 6: Summary of Recommendations 33

Section 7: Outcomes 34

Overview 34

Specific Considerations and Recommendations 34

Section 7: Summary of Recommendations 36

Section 8: Bias, Confounding, and Effect Modifiers or Subgroup Effects 37

Overview 37

Specific Considerations and Recommendations 37

Section 8: Summary of Recommendations 39

Section 9: Statistical Methods 39

Overview 39

Specific Considerations and Recommendations 40

Section 9: Summary of Recommendations 41

Section 10: Study Findings 41

Overview 41

Specific Considerations and Recommendations 42

Section 10: Summary of Recommendations 42

Section 11: Interpretation and Generalizability 43

Overview 43

Specific Considerations and Recommendations 43

Section 11: Summary of Recommendations 44

Section 12: Limitations 44

Overview 44

Specific Considerations and Recommendations 45

Section 12: Summary of Recommendations 45

Forward-Looking Statement and Conclusion 45
References 47
Appendix 1: Methods 51
Appendix 2: Documents Reviewed for Candidate Recommendations on RWE or RWD Reporting for Expert Survey 57
Appendix 3: Recommendation Checklist 59

List of Tables

Table 1: Categories of Data Extraction Matrices 52
Table 2: Results From First Methods Expert Panel Consensus Survey 54
Table 3: Recommendation Checklist 59

Abbreviations

CIHI	Canadian Institute for Health Information
COS	core outcome set
HTA	Health Technology Assessment
ICES	Institute for Clinical Evaluation Sciences
INESSS	Institut national d'excellence en santé et en services sociaux
RCT	randomized controlled trial
RWD	real-world data
RWE	real-world evidence
WG	working group

Summary

Real-world evidence (RWE) is evidence on the use, safety, effectiveness, and cost of health technologies that is derived from real-world data (RWD).[2,3] Regulators, Health Technology Assessment (HTA) agencies, and other stakeholders have recognized the necessity of incorporating high-quality RWE to help address evidence gaps for decision-making.[3-6] However, as the volume and types of RWE have rapidly expanded, there is a need to promote standardization in the reporting of HTA and regulatory studies involving RWE.[3]

The variety and complexity of RWD sources, study designs, and analytical methods make the evaluation of RWE studies challenging. Therefore, to optimize the utility and transparency of RWE studies to HTA and regulatory bodies, it is important to develop a set of common principles and core standards for reporting. Recent global initiatives have focused on developing tools to improve reporting, transparency, and reproducibility of RWE studies. However, there is a need for further guidance and standards to ensure that adequate information is provided to allow for thorough appraisal of RWE studies by regulators and HTA bodies. Additionally, other reporting guidance documents have largely aimed to establish reporting standards for either HTA or regulatory uses, but not both.[7] To address these needs, this document serves as comprehensive, credible, and fit-for-purpose reporting guidance that aims to harmonize current RWE principles for Canadian HTA agencies and regulators while maintaining alignment with international standards. The recommendations in this guidance focus on transparent reporting, which will facilitate HTA and regulatory appraisal of RWE.

What Is RWE?

Randomized controlled trials (RCTs) are the gold standard for establishing the efficacy and safety of health technologies.[8,9] However, trials often produce results for specific target populations and settings within controlled environments, thus limiting the generalizability of results to patients in real-world settings. Additionally, in some circumstances, such as the evaluation of drugs for rare diseases, trials that are sufficiently powered to detect clinically meaningful treatment effects for important clinical outcomes are not always feasible. RWE can potentially provide more generalizable evidence that fills knowledge gaps on the effectiveness, safety, and cost of drugs, medical devices, and clinical interventions.

RWD are data relating to patient status and/or the delivery of health care collected from a variety of sources, and can include electronic medical records, clinical and disease registries, and administrative databases.[2,10,11] RWD can also be drawn from other prospective sources, including pragmatic and hybrid trials. RWD can provide information about medical history, demographics, socioeconomic factors, health behaviours, experiences, clinical and functional outcomes, resource use, and costs.

RWE is defined as the evidence surrounding the usage and potential benefits or risks of a medical product derived from analysis of RWD.[2] RWE stemming from RWD can offer certain advantages over clinical trial evidence, such as the inclusion of patients who are underrepresented in trials, like children or older adults, patients from diverse ethnic groups, underserved and understudied populations, or patients with a high burden of multimorbidity. In addition, RWE about these populations can leverage expanded sample sizes and longer follow-up periods that may not be feasible in clinical trials to inform decision-making on

use, effectiveness, safety, and patient experience with health technologies. RWE can also offer insights into health care providers' and patients' or caregivers' perspectives on issues related to accessibility, acceptability, preferences, and ease of use of health technologies. There is potential to leverage RWE across phases of the health technology development life cycle; for example, RWD can be used to estimate the number of patients with rare diseases who may benefit from a new health technology, or to provide an assessment of off-label efficacy of medications and devices. Regulators and industry partners have long relied on RWE in pharmacovigilance and adverse event monitoring and reporting.

However, RWE has inherent limitations and is not necessarily appropriate to provide evidence in all scenarios. Given the limitations with causal inference from RWE studies, RWE should not replace clinical trial evidence, but rather is expected to supplement existing trial evidence as part of the broader body of evidence for decision-making. For instance, RWE is often subject to bias and confounding. Issues with RWE often include nonrandom assignment to treatment, and unblinded ascertainment of outcomes that may not be adjudicated and verified with the same degree of rigour as in clinical trials. Further, the generation of RWE requires many complex decisions and can vary largely in quality;[3] thus, reaching appropriate conclusions from RWE requires transparent reporting and careful interpretation of the RWD source, study design, and methods. Clear standards are needed to guide the reporting of RWE for decision-making to ensure adequate understanding of these complex decisions.[3]

Overall Purpose and Main Objectives

The overall purpose of this guidance is to promote standardization in the reporting of RWE studies for those undertaking and submitting RWE studies to support decision-making in Canada.

The development of this guidance was founded on the following main objectives:

- to ensure that regulators and HTA agencies have sufficient information to evaluate a study for its appropriateness of use for decision-making
- to provide core reporting standards for RWE studies that align with global standards
- to prioritize transparency in reporting while accounting for practical challenges related to RWD and RWE.

Specific Objectives

The specific objectives were to:

- identify existing global guidance on principles and standards for reporting on RWE studies to create an initial draft of Canadian RWE reporting standards that align with international standards
- establish consensus on items to be included in the core reporting standards for Canadian RWE studies through engagement with national and international experts in RWD and RWE.

About This Guidance

Standards for generating and reporting RWE have already been developed on a global level (refer to Appendix 2); as such, this document was created to best align with international standards while providing

specific consideration of the Canadian context. The aim is for all submitted RWE to provide detailed reporting that is relevant and useful for Canadian HTA agencies and regulators. Importantly, many of the components of this guidance focus on ensuring the highest level of transparency possible in the reporting of RWE studies.

The present guidance aims to ensure that each RWE study will provide Canadian regulators and HTA agencies with the information they require to appraise the study, to determine if and how the evidence should be used to inform decision-making. Due to the complexity of the use of RWD and the substantial reporting that is required, this guidance is not intended to educate or train readers on how to generate RWE; it is written for an audience that is technically versed in RWD and RWE methods.

This document is a critical first step for strong foundational guidance on the overall reporting of RWE. However, this guidance will be in a living document that will require updates, revisions, and extensions over time. This guidance document allows sufficient flexibility for uses to accommodate the heterogeneous nature of RWE and its rapid evolution, while ensuring that studies are sufficiently detailed and transparent to facilitate regulatory and HTA appraisal and decision-making. Lastly, this document does not provide guidance as to *when* or *why* RWE should be used (e.g., whether an RWE study is appropriate for a particular research question or when to use RWE), but rather how to transparently report RWE.

Implementation Considerations

There are important considerations for operationalizing the recommendations provided in this document. First, there are diverse RWD sources, RWE designs, and uses of RWE. This guidance was written in a manner that allows for flexibility in its use for a variety of RWE applications. As such, some recommendations in this document will not apply to all RWE studies. For example, utilization or burden of disease studies may not require all components described in this guidance. Second, we recognize that some therapeutic areas, such as rare diseases or medical devices, present unique challenges in the generation of RWE (e.g., identifying comparator groups). The recommendations in this document have been designed to allow sufficient flexibility to address and describe these challenges in studies.

We anticipate that this document will be updated periodically to incorporate lessons learned as the science evolves. Future efforts can address challenges observed in the operationalization of recommendations for reporting RWE and explore potential extensions of this document as the field of RWE evolves.

Background and Methods

This guidance was developed through an iterative process with the support of Canadian and international RWE experts and stakeholders. Full details of how this guidance was developed are reported in Appendix 1. In brief, a state of knowledge report created by INESSS[5] and an environmental peer-reviewed literature scan by CADTH on the use of RWE and RWD to support decision-making in drug assessments[12] were leveraged and expanded to identify relevant documents surrounding international RWE guidance, including systematic reviews, reporting guidelines, and policy statements. Additional documents were identified using a citation-review method and expert consultation. The resulting set of documents was reviewed to develop candidate reporting recommendations for expert review; as such, all recommendations were based on existing

guidance and reporting standards. In total, 37 documents were reviewed (refer to Appendix 2 for the full list). Recommendations on the reporting and conduct of RWE from all identified documents were independently extracted by 2 investigators; a third investigator reviewed the extracted data for accuracy.

All recommendations across documents were organized into a matrix categorized by type of recommendation (i.e., reporting versus methodological considerations) and study component (e.g., exposures). A total of 200 candidate recommendations were included in a questionnaire developed by the methodological authorship team that was shared among the group of 15 Canadian and international experts. Experts were asked if each candidate recommendation should be included in guidance for standards on the reporting of all RWE studies intended for any health technology or regulatory use in Canada. A recommendation was included or excluded based on whether there was 70% or greater consensus on its importance by the experts. Recommendations that did not achieve 70% or greater consensus were discussed in a large group meeting and revised as needed or excluded. Experts had a chance to flag recommendations they wished to discuss even if consensus was reached, and this was done for 1 recommendation. A draft report was shared with the experts for review and feedback, which was subsequently collated and incorporated by the Methods Authorship Team. Outstanding points of disagreement were discussed at a second in-person meeting, and consensus was achieved using the same process previously described.

Throughout its development, the guidance document was reviewed for alignment with current international standards and its suitability for the Canadian context for health technologies. Additionally, a robust stakeholder consultation process and public feedback period was facilitated to ensure engagement with members of the Canadian health technology ecosystem. A draft report was posted on the CADTH website for public and stakeholder feedback for 8 weeks. The Methods Authorship Team and Leadership Review Team participated in multiple in-person and virtual events, and leveraged established networks to provide multiple opportunities for additional feedback to be submitted. The Methods Authorship Team and Leadership Review Team collaboratively reviewed the feedback and grouped comments into either general themes that could be applied throughout the document (e.g., consistency of language) or RWE reporting–specific feedback. Major revisions to the document to address RWE methods–specific feedback were presented to the Expert Methods Panel for final approval and included or excluded based on whether there was 70% or greater consensus from the Panel. Items not reaching 70% or greater consensus or flagged by a panel member were discussed in an Expert Methods Panel meeting and revised or excluded thereafter, as needed. The Expert Methods Panel members then reviewed the document after incorporation of major revisions and changes.

Overview and Structure

The guidance is reported in 12 sections. Each section presents an overview and a narrative of the recommendations in detail. A summary list of recommendations completes each section. Some sections may have overlapping concepts, but each has a specific goal and purpose. The summary lists of recommendations and the checklist to be used when preparing for submission to regulatory and HTA

organizations (Appendix 3) are not intended to replace a careful review of the text, which contains critical information needed to develop adequate reporting.

Guidance is provided for the reporting of study components as follows:

1. Research Questions and Study Design
2. Setting and Context
3. Data Specifications, Access, Cleaning Methods, and Linkage
4. Data Sources, Data Dictionary, and Variables
5. Participants
6. Exposure Definitions and Comparators
7. Outcomes
8. Bias, Confounding, and Effect Modifiers or Subgroup Effects
9. Statistical Methods
10. Study Findings
11. Interpretation and Generalizability
12. Limitations

Section 1: Research Questions and Study Design

Overview

Full transparency in reporting research or study questions and study designs allows for a straightforward interpretation of study design decisions, and also serves to facilitate the reproducibility of RWE. RWE may aim to answer a wide array of questions (e.g., safety, effectiveness, or uptake of medications) and therefore can leverage a variety of study designs. Similarly, the same RWD source can be leveraged for a variety of designs when producing RWE, including traditional observational studies (e.g., cohort studies) as well as pragmatic trials.[13,14] Study designs are not compared in this guidance in terms of their strengths and limitations; rather, we advocate for transparent reporting and justification of the design choices to facilitate a robust assessment, regardless of the specific study design used. Detailed reporting of the study objectives and study design facilitate reviewers' ability to understand, interpret, and appraise the design and methods.

Importantly, asking causal research questions from RWE studies can be challenging. However, causal inference is sometimes the goal of an RWE study, especially when related to effectiveness claims. Causal inference from RWE is possible when appropriate design choices are made, rigorous methods and fit-for-purpose data are used, and assumptions are met. Investigators are directed to additional reading to learn more about causal inference methods.[15]

Specific Considerations and Recommendations

Study Aim and Research Question

The aim and study question must be clearly reported. An aim is the overarching goal of the research study, and the study question is the specific intent of the study. The study question should be phrased by using the Population, Intervention, Comparator, Outcome, Timing, and Setting (PICOTS) template.[16] The PICOTS template can be adapted for varying study designs depending on the research question being asked. Each component of PICOTS should be reported in a detailed manner and in line with any relevant literature. To support the rationale for the study aim and study question, there should be a broad review of the relevant literature to provide pertinent background information and outline current gaps in knowledge.

Study Design

The study design (or multiple designs, if used) should be reported. The rationale for the choice of design should be supported by relevant literature. Although detailed later in the guidance, primary and secondary outcomes and the main measure(s) of effect (e.g., hazard ratios) should also be specified, as they are likely influenced by the selection of the study design. Depending on the study design(s) used, other important components that must be reported include: design descriptions for study arms (e.g., parallel or crossover); allocation ratios between study arms; and, if matching is implemented, clear reporting of the overall allocation ratio and matching criteria (e.g., 1-to-1 hard-matching based on age and sex). It is suggested that reporting be aligned with established standards of reporting for the type of study design employed.[17-21]

Studies with causal research questions should consider modern causal inference frameworks, such as target trial emulation, to guide their study design and methods.[15,22,23] Target trial emulation is a framework wherein the investigator specifies how each component of their study design (e.g., eligibility criteria, interventions or treatment strategies, outcome, follow-up, causal contrast [comparison], and statistical analysis) could be implemented in analogous ways to a randomized trial, so that critical biases (i.e., immortal time and selection bias) can be avoided and causal effects can be estimated.[24] Interested readers can refer to detailed texts on the target trial causal inference framework for more information.[24]

To improve transparency and ensure the interpretability, the use of study design diagrams is suggested.[25] These diagrams illustrate important components of the study design, including observation windows, exposure windows, covariate measurement windows, washout windows, and lag windows. If a study diagram is not used, all of these components must be fully described in the text. If diagrams are leveraged, it is suggested that best reporting practices be used as recommended by current standards.[25]

Other Recommendations and Additional Transparency

Development and registration of an a priori protocol before conducting the study is strongly recommended, particularly when the RWE is intended to provide confirmatory evidence of effectiveness. The protocol should also be discussed early with regulators and any relevant HTA agency if the planned study is intended to support a submission. Investigators may consider using a standardized protocol template to develop their study protocol, such as the HARmonized Protocol Template to Enhance Reproducibility (HARPER).[26] The protocol should be registered to a permanent platform that assigns a unique study identifier and is

maintained by a third party. Examples of such platforms with these qualities at the time of development of this guidance are: the Real-World Evidence Registry,[27] ClinicalTrials.gov (Observational Study Type specification),[28] and the European Union electronic Register of Post-Authorisation Studies (EU PAS Register).[29] If a study protocol was developed, it should be referenced in the initial reporting of the study design, including a citation or reference for the protocol and the registration number.[25] Any deviations from the protocol must be detailed, including each change, why the change was enacted (with a justification), and when this change occurred in the study process. It is important to note that a priori analyses (i.e., those outlined in a protocol) are preferred to post hoc analyses. Reporting of any research ethics approvals (or equivalent) or an ethics committee approval waiver, with reference numbers, is required when the content of the study is considered research. Note that the definitions of research and human subjects research can differ by the institution and research ethics board responsible at the location(s) where the work was conducted.

It is suggested to include a description of each team member involved in the study, specifying their role, organizational affiliation, education, title, and experience. It is strongly recommended that each team member disclose any potential or actual conflicts of interest, which may be financial or nonfinancial and direct or indirect. Inclusion of patient partners throughout all stages of the research process is recommended, and their role and degree of involvement should be clearly described.[30,31] In addition, highlighting which team members have hands-on experience and knowledge of the data source may increase the reviewers' confidence in the appropriate use of the data. Team inclusion should align with the International Committee of Medical Journal Editors (ICMJE) standards,[32] recognizing that some team members may not be authors based on differing rules and regulations of organizations. Lastly, reporting of study governance, especially if multiple partners are involved, is recommended to allow for full transparency of the study structure and execution. Study governance reporting must include all sources of funding and potential conflicts of interest for external groups involved, if applicable, and must specify who had decision-making power and final approval. Decision-making power includes but is not limited to input on any aspect of the methods, design, or interpretation of the results.

Section 1: Summary of Recommendations

- Report a clearly stated aim and study question.
- Report the overall study design.
- Provide a rationale for the choice of study design.
- Provide a relevant review of the literature to evaluate pertinent information and gaps in knowledge.
- Describe key elements of the study design (e.g., matching).
- Consider the use of study diagrams to illustrate key aspects of the study design.
- Strongly recommend developing and referencing an a priori protocol.
- Describe all study team members including the role of patient partners and any conflicts of interest.
- Describe the study governance structure, especially who was responsible for final decision-making.
- Report any research ethics approvals (or equivalent).
- Disclose sources of funding.

Section 2: Setting and Context

Overview

There is a potential to leverage RWD and evidence from multiple jurisdictions, particularly for the study of rare diseases. Moreover, replicating study findings using another data source can provide important information and enable validation or triangulation of results, depending on the heterogeneity of databases and populations. However, with any RWD source, it is essential that all reporting include detailed information on the study setting and context (i.e., health system factors like universal health care, private payer), even if the data come from a Canadian setting. This reporting is especially critical for datasets that are not regularly used by various research groups or are novel (e.g., single system electronic health record data, novel registries, or patient support programs). Even when well-known or common datasets (e.g., data sourced from Statistics Canada, the Canadian Institute for Health Information (CIHI), or ICES databases) are used, detailed reporting should be provided, as data sources change over time and the degree of detail and types of data accessed can differ by study. Non-Canadian data can be an acceptable source of RWE but must have important components reported in order for reviewers to understand the RWD and RWE's strengths, limitations, generalizability, and transferability to the Canadian context.[33]

Whether an RWD source is acceptable may depend on whether the data are fit for purpose with regard to the specific question being asked, the level of data access, current version of the database, and other factors including data completeness and accuracy. Thus, a list of suggested or acceptable RWD sources cannot be provided. The reporting components in this section can help study reviewers understand the implications of any RWD source used, including the context in which the RWD was collected and the data source's strengths and limitations. The following reporting components relate to several subsequent sections in this document (e.g., Participants); therefore, this section can be referenced for reporting and justification of components in these respective sections.

Specific Considerations and Recommendations

All Data Sources

The setting in which the data are collected must be clearly and fully described. The setting includes the geographic location(s), health care system context, and time periods in which data were available. Reporting on the health care system context should include the overall health care system structure (e.g., universal coverage), care models (type and number of care providers), sectors (e.g., primary care, inpatient settings, specialist care, nursing homes), population size (including the proportion of the population included in the final study), and payment structures (e.g., capitation, fee-for-service). All relevant study period dates should be described, including periods of recruitment, exposure, follow-up, and data collection, as applicable. Missing data components that arise due to the data setting that are pertinent to the study question of interest should be clearly specified. For example, if the study question surrounds medication adherence and drug exposure data are not available during hospitalization periods, this should be acknowledged.

Non-Canadian Data Sources

Non-Canadian sources may be acceptable sources of RWE, but given the importance of generalizability, studies leveraging non-Canadian data must have important components reported. Foremost, the rationale for why Canadian data were not utilized should be described; justification for use of the data source and its alignment with the study aims and study questions must be articulated. In addition, an explanation of how study setting factors might affect the generalizability of the results to the population in Canada must be provided. Participant demographics (e.g., age, sex, gender, and race or ethnicity), as well as the incidence and prevalence of the disease, confounders, and effect modifiers (if applicable) should be reported in the justification for transferring non-Canadian data to the Canadian context. Background information about the health care system, including methods of diagnosis, diagnostic criteria, standard patterns of treatment for the disease(s) of interest, and the degree to which such information is collected in the proposed data sources should be described. Furthermore, a description of prescribing and utilization practices, including approved indications, formulations, and doses for the treatment(s) of interest in the non-Canadian setting should be included.

Information on the market availability and major changes in use of the intervention and comparators of interest throughout the study period in non-Canadian settings (e.g., regulatory approval dates, formulary restrictions, and major health policy or health care changes [e.g., the COVID-19 pandemic]) should be reported, highlighting if and how they differ from the Canadian market.

Section 2: Summary of Recommendations

- Describe important information to contextualize the data source, including:
  - type of care setting
  - geographical location.
- Describe all relevant study period dates, including periods of recruitment, exposure, follow-up, and data collection.
- Clearly identify missing data components in the data collection.
- For studies that propose the use of a data source from a country other than Canada, provide:
  - a rationale for selecting the data source
  - an explanation of how all the factors might affect the generalizability of the study results to the population in Canada
  - background information about the health care system
  - a description of prescribing and utilization practices
  - information on the use and market availability of the intervention and comparators of interest throughout the study period.

Section 3: Data Specifications: Access, Cleaning Methods, and Linkage

Overview

Utilization of RWD often requires many steps related to the access, cleaning, and linkage of data sources before analysis begins. Detailed guidance on the conduct of data quality control and provenance is available.[34,35] Reporting on data provenance is important to ensure credibility of the data leveraged and full transparency of data specifications. The current landscape of data access is complex, with various nuances related to data ownership, privacy regulations, and intellectual property that may be barriers to full reporting of data specifications. Given the importance of transparency with RWE, it is suggested that all specifications be reported to the most detailed extent possible. The inability to report any of these components of data provenance may limit the interpretability of the study and should be highlighted as 1 or more limitation(s), as appropriate.

Specific Considerations and Recommendations

Data Access

Describe the extent to which the investigators had access to the data. Briefly describe data provenance (the origin of the data, data custodians, data governance, and major transformations before investigator access).[35] Data ownership and processes for access must be described, including whether a data vendor or organization was used, whether other researchers can access the data, and whether costs are associated with data access in general (information that helps to understand the degree to which replicability studies could be conducted). These statements also apply to registry data access. Clearly describe any difference(s) between the source data and the data used for the analysis (e.g., sampling, information suppression).

Data Cleaning

Provide information on the data-cleaning methods used in the study. Describe transformations to the data fields to handle missing or out-of-range values, duplicate records, or logical inconsistencies. Provide code with annotation — or reference previously-published code — to identify key operational and design parameters related to data-cleaning algorithms. If these components cannot be reported for the entire data source, it is recommended to report these steps for the analytical study data at a minimum. Report whether data were organized by a Common Data Model structure (i.e., a uniform set of metadata or variables, such as the Observational Medical Outcomes Partnership [OMOP] Common Data Model[36] or the CIHI Reference Data Model).[37] Best practices for data cleaning of unstructured data (e.g., free-text clinical notes) are evolving, given the more recent use of these RWD. Studies leveraging unstructured data should follow up-to-date guidance on best practices for their use.[35,38]

Data Quality

Characteristics of data quality must be reported, including data completeness, validity of any data-cleaning algorithm(s), data extraction, and transformation processes. Data completeness refers to the percentage of records without missing data at a given time point.[35] Describe established routine data quality checks

and any internal and external audits that were conducted. Describe the extent of missing or out-of-range values, logical inconsistencies, and reports of persistence (the degree to which data values are consistently accessible over time).[35] Any variability between data sources and the impact of changes over time in the data should be reported (e.g., pre- versus post-onset of the COVID-19 pandemic).

Data Linkage

State whether the study included person-level, institutional-level, or other levels of linkage across databases. Report if consent was required for database linkage and, if so, how it was attained. Describe the methods of linkage, including whether the linkage was deterministic or probabilistic,[39] which variables were used for linkage, and which entity performed the linkage (e.g., data provider versus study analyst). The performance characteristics of data linkage must be described (e.g., proportion not linked or matched, and changes in linkage performance over time). If available, the number of individuals with linked data at each stage should be reported if a multistage approach was used, to better understand how representative the final study data are of the population of interest.

Other Recommendations

Any methods used for primary data collection should be clearly described, if applicable. For example, if questionnaires or surveys are involved, complete copies of data collection forms (including skip patterns) should be provided. If the study or registry required individual informed consent for recording personal data (the registry's primary purpose), provide the consent document (document file format); or, if regulations exist for data management in the absence of informed consent, describe the relevant regulation(s) or permission(s) received.

Section 3: Summary of Recommendations

- Describe the extent to which the investigators had access to the database population used to create the study population and major aspects of data provenance.
- Provide information on the data-cleaning methods used in the study. Share any data-cleaning code leveraged. If not provided, justify.
- Report whether data were organized by a common data model structure.
- Describe the usage of data and consent for data sharing. Provide consent documents, if relevant.
- Describe data collection methods.
- Report the quality of the data and relevant metrics used to assess the data quality.
- Describe any variability between data sources and the impact of changes over time in the data.
- Describe if any data linkage was conducted and the methods used for the linkage.
- Report who (e.g., which organization) performed the data linkage, if applicable.
- Describe the performance characteristics of the data linkage and the number of individuals linked at each stage of linkage.

Section 4: Data Sources, Data Dictionary, and Variables

Overview

In addition to transparent reporting of data access, cleaning methods, and linkage, clear reporting of data sources used to measure all variables is equally critical to understand the study methodology and facilitate reproducibility. Different geographical locations and settings may result in varying data availability, continuity, and completeness; therefore, characteristics of the health setting and context of data collection must be described. Importantly, RWD are commonly accessed through public, not-for-profit, and private data vendors and custodians for research purposes (e.g., CIHI, Statistics Canada, ICES, registries). Therefore, the names, dates, and/or version numbers of data extracted for research use, along with the dates and additional search and/or extraction criteria applied to create subsets of data, must be clearly and fully described where possible. For each variable of interest, the data sources, methods of measurement, and validity, as available, are needed to provide insight to applicability to a real-world Canadian context.

Details on the methods used to define study variables are critical for a study leveraging RWD. Major study variables include the exposure, outcome, potential confounders (covariates), and effect measure modifiers. In particular, the lookback windows and any time-varying definitions used in the measurement of these variables are critical to report. Detailed information on these variables must be included within a data dictionary.

Specific Considerations and Recommendations

Data Sources and Context

Indicate all sources of data being used in the study, including how they were obtained, and provide justification for the data selection. Report the specific version of the database used and the date of the last update of the database, if available. Describe characteristics of the health setting with mention of the geographical location, type of setting, and context of data collection. This information is particularly important for research involving multiple jurisdictions where the availability of data, such as prescription records, may differ. Data continuity, comparability, and completeness must be clearly described across data sources. Include descriptions of how and why gaps in data coverage may occur. For investigators or analysts using administrative claims data or registries, reporting the data completeness (i.e., continuity of coverage) is important, as individuals often enrol and disenrol in different health plans in relation to changes in employment or other life circumstances. Any variations among source data (e.g., intrapatient or interprovincial differences in data availability) should be documented. Specify the source(s) of data for each major variable of interest in the study and briefly discuss whether the data source(s) can validly measure the study population(s), exposure(s), outcome(s), and key covariates. For registries, investigators may consider using tools like the Registry Evaluation and Quality Standards Tool (REQueST) to provide context on the quality and content of the registry data.[40]

Data Extraction

It is imperative to report the names, dates, and/or version numbers of all contributing sources of data. If extraction criteria were applied to create a subset of data used for the research, detailed descriptions of the criteria are needed as a means of understanding cohort development. Extraction criteria should also include calendar date ranges, as data continuity may also be affected by time, particularly for commercial data sources where participants may change from year to year. Providing explicit extraction criteria and date ranges facilitates reproducibility and adds additional verification of the process through which the final study population was reached. If this information is not readily available, a request for additional information from the data vendor should be considered and included, if acquired.

Data Dictionary

Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable.[41] Include information on how types of data for variables were collected; timing of capture, including the lookback window; and the source (e.g., clinical diagnoses, tests, procedures, prescriptions). It is important that naming and variable definitions remain as consistent as possible. Report how all variables were coded, recorded, or collected, as well as validation of the quality of the variable, if known. Report important variables that were not available in the data source and justify why they were not included. It is important to recognize that a lack of certain information and data may limit the ability for reviewers to assess the use and appropriateness of these variables. If multiple data sources or multiple versions of the same data source were used, report any differences in how data were coded, recorded, or collected between sources, and how those differences were reconciled or addressed. For primary data collection, specify any quality assurance processes that were in place (including training or blinded review). The data dictionary should also contain any deviations in the study from the a priori protocol; specify how adaptations were allowed and recorded with the dates of each amendment.

Types of Variables and Measurement

For each major study variable of interest, specify the lookback window used to ascertain variables (e.g., in the 365 days before the date of first exposure). More detailed considerations and recommendations on the measurement and reporting of exposures and outcomes are provided in sections 6 and 7, respectively. In general, confounders and effect modifiers should be defined before the exposure to avoid adjustment for causal intermediaries (factors on the pathway between the exposure and outcome). The description of variable measurement should also include whether any variable could be time-varying, with details on how the variable could change over time and when it was redefined in relation to time-varying exposures. Reporting requirements for predictive modelling and mediation analyses are outside of the scope of this document, but reporting of predictor variables and mediators (intermediaries) should be similar to reporting for potential confounders and effect modifiers.

Section 4: Summary of Recommendations

- Provide and describe all data sources, including the specific version and date of the last update of the database.

- Describe the characteristics of the health setting and the context of data collection.
- Describe details of data continuity and completeness.
- Include the names, dates, and/or version numbers of when data were extracted for research use by the data vendor or organization.
- Include the search and/or extraction criteria applied if the source data are a subset of the data from the vendor or organization, and provide calendar date ranges.
- Provide source(s) of data for each variable of interest.
- Describe how variables of interest were measured and if they have been adjudicated or validated in the population of interest.
- Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable.
- Specify definitions and lookback windows for all variables.
- Report whether any variables could be time-varying (e.g., how the variable could change over time and when it was redefined in relation to time-varying exposures).
- Report important variables that could not be captured and their anticipated impact on study results.
- Provide information on deviations from the a priori protocol in variable measurement.

Section 5: Participants

Overview

The details of participant selection are essential for understanding how generalizable the study population is to the real-world target population, and to address potential issues related to selection bias. All methods and decisions that led to the final, analyzed study population (e.g., random sampling from the source population, inclusion and exclusion criteria) should be described in a stepwise manner, with definitions provided (e.g., exposure groups, cases, controls). Ideally, these steps should be described in a figure (e.g., an "exclusion flow" figure that presents included and excluded patients in a stepwise fashion from the original source to the final analytical sample).[42] Investigators must clearly describe all inclusion and exclusion criteria used to identify the study population along with detailed justification for each exclusion criterion.

A detailed description of study participant characteristics is also critical for assessing potential confounders or bias, evaluating the safety and effectiveness of drugs or treatments of interest, and determining generalizability of the findings. Accurate reporting of these characteristics is needed to determine who may benefit from a certain treatment and, conversely, who may be at risk of harm. Reporting the numbers of participants at each stage of the study should be accompanied by reasons for losses to follow-up or nonparticipation, if available. It is recommended to use a visual aid or figure to represent the reported number of participants at each stage of the study.[42] If conducted, statistical comparisons of participant characteristics between treatment or exposure groups must be described along with a description of the extent and handling of missing data.

Specific Considerations and Recommendations

Inclusion Criteria

Describe all inclusion criteria and the order in which these criteria were applied to identify the study population. Specify any enrolment requirements (e.g., participants who contributed to a data source for a defined period) that were required for inclusion or why they were not necessary. Indicate whether participants were entered into the study population only once, or if multiple entries were permitted. Report if enrolment gaps were allowed before inclusion or during the follow-up period. Specify if a temporal window (e.g., lookback window) was used to assess inclusion and exclusion criteria.

Discuss and explain how the selected study population compares to the target population (i.e., real-world patients). Specifically, describe the study population characteristics, including age range, sex, gender, comorbidities, medications, and any other important factors in comparison to the target population. If a non-Canadian study population is used, investigators should refer to the reporting recommendations for non-Canadian data sources in section 2 (Setting and Context). In addition, this section should discuss the study's inclusion and representation of patients by sex or gender, race or ethnicity, and other characteristics important to consider for diversity — according to up-to-date guidance —[43] and should emphasize how historically underrepresented groups in research are included, to the extent that is possible.[43] In alignment with other international standards and guidance (Appendix 2), all codes and/or algorithms (e.g., drug, diagnosis, procedure, lab codes) used to define the inclusion and exclusion criteria should be specified where possible. The study must acknowledge when this information cannot be provided for inclusion and exclusion variables. In the event that codes or algorithms used for study variables are proprietary and cannot be shared, it is strongly suggested that investigators provide at least an overview of the concepts or steps involved in these algorithms. If validation studies of the codes and algorithms used for inclusion or exclusion were previously conducted, cite these. If validation was conducted for this study but not published elsewhere, provide detailed methods and results of the validation study.

Exclusion Criteria

If a particular group of patients was excluded from the study, investigators should justify this approach, providing a detailed explanation of the exclusion, the order of exclusion criteria applied, and any resulting limitations in the interpretation of the findings. As listed above, all codes or algorithms used to define inclusion and exclusion criteria should be reported, where possible, as well as any temporal window used to assess these criteria, along with the calendar date range.

Cohort Studies

Further to the inclusion and exclusion criteria outlined above, there are items pertaining specifically to cohort participant selection that should be reported. For study cohort design, indicate whether a new-user,[44] prevalent-user, or other type of cohort study design was used. A new user of a medication may refer to a person identified at the first use of a medication after some established period without prior use (sometimes referred to as a "washout period"; e.g., 180 days without use of the intervention or drug of interest). Where possible, a new-user design is preferred because follow-up for all persons begins at the same time point in the treatment course. In contrast, the prevalent-user design often begins follow-up for persons at different

time points in their treatment course, potentially resulting in issues such as depletion of susceptibles (e.g., some persons have already experienced outcomes of interest and discontinued therapy before study accrual begins).[44] If a new-user design was chosen, specify the lookback window used to ensure participants were new users of the treatment(s) of interest. If a new-user design was not used, justify the choice of cohort design.

If it is a comparative analysis, data on the number of participants in the exposure group(s) at each stage of cohort development are needed to determine how the final cohort was established. These data include the number of participants before the application of exclusion criteria and at each exclusionary step, and the analyzed study population. For matched cohorts, matching criteria should be described, if applicable, in addition to when the follow-up period began (i.e., the index date, time-zero, or cohort entry date). Specify when follow-up of a participant stopped, including reasons for censoring or whether follow-up ceased at first outcome/event. If censoring was applied, report the number of participants in each exposure group that were censored due to each censoring criterion. We recommend this entire process be represented with a figure to clearly communicate the sample size at each step.

For prospective cohort studies, describe the cohort recruitment process and discuss whether the cohort is reasonably representative of the target population. Acknowledge whether some patient groups may not be represented and, if so, how non-inclusion of these groups may limit the external generalizability of research findings.

Case-Control and Case-Crossover Studies

As for cohort studies, there are items pertaining to participants and study design in case-control and case-crossover studies that are necessary to report. For example, if the case-control study is nested, describe the cohort or source population from which it is derived as discussed previously. Discuss the methods of case ascertainment and control selection. Describe and justify the methods for the selection of controls, including matching criteria, any sampling methods, the number of controls for each case, use of calipers, and hard-matched covariates between cases and controls (e.g., sex).

Study Participants

Provide the numbers of participants at each stage of the study (e.g., participants potentially eligible, examined for eligibility, confirmed eligible, included in the study, completed follow-up, and analyzed).[20,21] Consider illustrating this information using a diagram to report the flow of participants throughout the study. At each stage, investigators should provide reasons for nonparticipation and/or exclusion. Investigators should provide a breakdown (e.g., in tabular format) showing the proportion of subjects lost to follow-up and/or excluded from the analysis, including the reasons why. Characteristics of study participants (e.g., demographic, clinical, social determinants of health, matching variables, exposures, and potential confounders) must be clearly presented, preferably through the use of tables (e.g., a "Table 1 of patient characteristics). For reporting of sex and gender, refer to the CIHR recommendations for the appropriate integration of sex and gender in research.[45] If any of these participant data are not available or feasible to obtain and report, explain why.

Disposition of Participants

Provide comparisons of participant characteristics by treatment or exposure groups. It will be important to indicate the number of participants with missing data for each variable (characteristic) of interest. Include the number of participants in each analysis conducted, the type of analysis (e.g., intention-to-treat) or whether a person's exposure group could change over follow-up (e.g., crossover). The impact of any exclusions from each analysis should be carefully assessed. Consider using standardized differences instead of hypothesis tests to compare patient characteristics between groups, as the results of hypothesis tests are largely dependent on the sample size.[46-48]

Section 5: Summary of Recommendations

- Provide inclusion criteria used to identify the study population.
- Justify exclusion criteria and how they may affect the overall interpretation of the research.
- Describe study population characteristics relative to the target population in Canada.
- Provide all codes or algorithms used to define the inclusion and exclusion criteria, where possible.
- Specify the time period (e.g., lookback window) over which inclusion and exclusion criteria were assessed.
- Recommendations for specific study designs:
  - For cohort studies, provide details leading to the analyzed cohort including definitions for exposure groups, cohort entry and end dates, matching criteria, and censoring/follow-up.
  - For prospective cohort studies, describe recruitment processes.
  - For case-control and case-crossover studies, provide details of case and control ascertainment, the source population for nested studies, sampling methods, and matching criteria.
- Report the numbers of participants at each stage of the study and reasons for nonparticipation. Consider illustrating this information using a flow diagram.
- Provide characteristics of study participants. If not available or feasible, explain why.
- Indicate missing data for each variable of interest.
- Compare treatment or exposure groups.
- Specify the number of participants included in each analysis and the analysis strategy (e.g., intention-to-treat) and provide details on the number or proportion of subjects excluded from each analysis and the reasons for exclusion.

Section 6: Exposure Definitions and Comparators

Overview

A critical reporting element of RWE studies is clear and justified definitions of the exposures and comparators used in the study.[25,49] The term "exposure" can refer to a host of treatments and factors,

including drugs, devices, or clinical interventions. This section focuses primarily on considerations for drug and medical device exposures. Defining the exposure enables reviewers to interpret the accuracy and completeness of an exposure definition. Defining the comparator allows reviewers to understand how the choice of a comparator controls for confounding by severity and indication. Exposure definitions should include information such as the data source(s) from which exposure information was obtained, limitations of the data source(s) to identify exposures (e.g., precise start and stop dates of exposures), and detailed requirements for the exposure definition (e.g., a requirement for a certain duration of use or multiple prescription fills for a medication). For a comparator group (or control period for self-controlled studies wherein subjects act as their own control), information such as the details of the comparator, justification of why this particular comparator was selected, and potential implications of comparator selection on study results should be provided. For studies that do not use any comparator(s), explain why. Additionally, specify if and how changes in exposure status (e.g., whether a person was within the intervention or comparator group) over study follow-up were permitted and measured. Finally, include any changes in the patterns of use of the exposure(s) or comparator(s) over time and how they may affect the study findings.

Specific Considerations and Recommendations

Exposures

Define the requirements for a patient to be considered exposed (e.g., single, multiple, or continuous exposures), as well as the start and stop windows for assessing exposures. For example, report whether a single prescription fill for the exposure medication of interest was sufficient for someone to be considered exposed, or whether they had to receive multiple prescription fills. Specify the data source(s) from which exposure information was obtained, including validity of the exposure measure(s) (e.g., the validity of prescription medication claims to measure true medication fills received by a patient), if available; and the description of limitations of this data source to capture the exposure(s) of interest. Describe any additional analyses used to assess the impact of changes to the exposure definition on the study findings. Specify the exposure-outcome risk window (e.g., whether events are attributed to current, prior, distant past/ever exposures, or cumulative drug exposures), and discuss how the window aligns with the known or suspected timing of the relationship between the exposure and outcome (e.g., instantaneous, delayed, dose-response).

Comparators

Specify the comparator used and provide justification for its use. When justifying the selected comparator group, consider areas of clinical equipoise such as the comparator's role in therapy (e.g., first-line versus second-line), access issues, and contraindications. Specifically, discuss the potential implications of the comparator group if it does not include:

- an active comparator
- a drug used to treat the same disease
- patients reasonably expected to have the same level of disease severity
- patients from the same time period as the exposed cohort.

If an external comparator was used, discuss how the study population and this external population compare and explicitly report any assumptions regarding the comparability of the external cohort. An external comparator is a comparator group, usually derived from RWD, that is compared to a group of patients that participated in a clinical trial. An external comparator might also be referred to as a "historical comparator," "external control," or "synthetic control." (Refer to current texts on implementation and best practices of external comparators in RWE.)[50-53] Similarly, if the comparator group is from the same individual but a time period before the exposure (e.g., in a self-controlled case series study design), explain whether there may be important differences in outcome risk between the exposed and unexposed time periods. Discuss whether formulary status or other medication access factors could impact the level of disease severity in the comparator group compared to the treatment group. If no comparator was used in the study, explain why.

Other Exposure and Comparator Considerations

Describe how exposure switching (i.e., changes in treatment over time) or dual exposures to the treatment and comparator were managed, if applicable. In addition, report any concomitant interventions (e.g., add-on therapies) and the extent to which they were used in each group. Discuss how changes in exposure status were handled during follow-up (e.g., whether exposed follow-up time was only when the participant was receiving the drug [as-treated], was ever on the drug [intent-to-treat], or other exposure definition). Sensitivity analyses to investigate the impact of potential exposure misclassification on study results are encouraged. Finally, discuss any major changes in patterns of use of the exposure and comparator over time (e.g., changes in access) and how they may impact the study findings. Report any methods that were used to adjust for the changes in treatment over time.

Section 6: Summary of Recommendations

- Define the requirements for the exposure definition (e.g., single, multiple, or continuous exposure) and relevant start and stop windows for assessing exposures.
- Specify the data source(s) from which exposure information was obtained, including validity and any limitations in exposure measurement.
- Specify the exposure-outcome risk window and discuss how it aligns with the known or anticipated relationship between the exposure and outcome timing.
- If no comparator was used, justify why not.
- Define the comparator group(s) (e.g., active comparator, historical comparator).
- Provide justification for the comparator used, including potential implications on the study findings.
- Discuss any changes in patterns of use of the exposure and comparator(s) over time and how they may affect the results. Report any methods used to adjust for these changes.
- Specify how adaptations to the intervention and/or comparator were permitted and recorded.

Section 7: Outcomes

Overview

The utility of RWE for decision-making around effectiveness relies heavily on whether the outcomes studied are relevant to the specific study question. It is also imperative that these outcomes are validly captured in the RWD used. Reporting on the selection and definitions of outcomes is therefore critical to the assessment of any RWE study. Here, the term "outcome" refers to the broad array of study end points and can encompass clinical events or other relevant measures for the disease and health technology being studied. This section must contain detailed information on:

- study outcomes and their definitions
- references on the validity of these outcome definitions (including the strength of association between any surrogate outcomes and clinical outcomes, if applicable and known)
- a discussion of the relevance of study outcomes to real-world practice
- considerations of outcome misclassification and the accuracy of outcome timing in relation to exposure to the treatment(s) of interest.

Specific Considerations and Recommendations

Outcome and End Point Definitions and Validity

Report which study outcomes were selected and specify whether each was a primary, secondary, or exploratory outcome. Exploratory outcomes are important events that are of clinical interest but are not assessed with the rigour needed to make conclusions, but are included to explore future hypotheses.[54] Report which outcomes were specified a priori versus which were post hoc. Report all changes to the planned protocol with reasoning. Provide rationale for why the study outcomes were selected (e.g., relevance to clinical practice, safety concern, patient or caregiver consultation, and data availability), citing evidence to support the rationale if available (e.g., outcome is clinically relevant based on previous study findings). Discuss any relevant outcomes or end points that were not studied with a justification as to why they were not included.

Specify the definitions used for all study outcomes.[55] If an outcome was assessed using objective criteria such as diagnostic codes, the definition provided should specify all codes or algorithms used to define outcomes, where possible. This definition should provide details on the exact codes used to identify the diagnosis, drug, procedure, or other event; whether inpatient and outpatient codes were used; and whether there were requirements for the coding position (e.g., primary diagnosis, secondary, any position), as applicable. The study must acknowledge when this information cannot be provided for outcome variables. In the event that codes or algorithms used for study variables are proprietary and cannot be shared, it is strongly recommended that investigators provide a high-level overview of the concepts or steps involved in these algorithms.

It is strongly recommended to report the validity of outcome measures. If validation studies of the codes or algorithms were conducted (i.e., studies that estimate the sensitivity or specificity of the code or algorithm),

reference these studies and report the performance characteristics and the population in which they were conducted. Discussion of the validity of outcomes should also consider whether the outcome timing could be assessed precisely in relation to the initiation and duration or discontinuation of the exposure(s). For example, outcome ascertainment of a myocardial infarction is likely relatively precise versus the onset of more insidious outcomes like dementia or cancer. If validation was conducted for the study outcome of interest and not published elsewhere (e.g., an internal study), provide detailed methods and results from this validation study, ideally with sensitivity, specificity, and positive predictive values of the outcome definition. If a sample of outcomes has been manually verified for validity, describe the sampling strategy and methods used to ascertain validity. If no validation studies of the outcome definition are available, justify why this outcome was used. Discuss any updates or changes to coding practices or versions for the outcomes across the study period (e.g., changes in International Classification of Diseases codes from the ninth to 10th edition), if applicable.

If an outcome is self-reported or observer-reported, specify whether a validated instrument was used and reference the validation studies.[56] If a validated instrument was not used to capture the outcome, then justify why not. Discuss whether the outcome or its measurement may be subject to clinical judgment (e.g., the outcome is a clinician's opinion on whether the patient's condition has improved). If applicable for the question, this section should also report whether outcome severity could be captured using the outcome definitions used. For example, in a study examining hospitalizations for COVID-19, investigators might discuss whether intensive care unit admission was able to be assessed. Discuss whether outcome severity might be different between treatment group(s), regardless of whether severity could be captured.

Adverse Event Studies

For studies that examine adverse drug events or reactions, specify whether events were assessed or validated on the individual case level (e.g., through record review by a specialist blinded to the exposure(s) under study, to try and rule out other more likely causes of the event). If so, specify the number of potential cases that lacked sufficient data to be classified as non-cases or definite cases (i.e., final status is "possible" or "uncertain").

Outcome Selection and Surrogate Outcomes

Discussing selected outcomes relative to their location on the causal pathway from the exposure is recommended. When available, clinical outcomes (e.g., major cardiovascular events) are preferred to surrogate outcomes (e.g., changes in laboratory values or biomarkers) as the primary outcomes. If a surrogate outcome is used, cite the strength of the relationship between the surrogate outcome and the relevant clinical outcome(s) (e.g., association between lowering of low-density lipoprotein values with reduction in myocardial infarction risk). A well-established and validated surrogate outcome should be selected, if available. If a surrogate outcome is not validated, justify its use (e.g., no other outcome was feasible or available for the study). At minimum, attempt to explore clinical outcomes as secondary end points if a surrogate outcome was used as the primary outcome. Consider the use of a core outcome set (COS) for standardized outcome reporting if 1 is available for the condition of interest under study.[57] Using a COS helps to improve consistency in reporting and comparability of studies for a particular condition while

reducing the risk of selective reporting bias.[53] A COS that was developed in conjunction with stakeholders, including patients or caregivers and/or patient organizations can help to ensure that outcomes being measured or reported for a particular condition are valued and patient-centred.

Other Outcome Considerations

Explicitly specify if outcomes were measured in the same manner for the treatment and comparator groups. Discuss whether the outcome may be differentially measured between the treatment and comparator groups; this discussion should include whether outcome data may be more likely to be missing or invalid for certain exposure groups. For example, patients receiving a medication subcutaneously at a physician's office may be more likely to have adverse events reported versus patients taking an oral medication at home. Likewise, patients with more multimorbidity may be seen more frequently for laboratory testing, and thus outcomes may be more likely to be captured in this group versus those with fewer comorbidities. Other important considerations that may result in differential outcome ascertainment between groups include differences in intercurrent clinical events during follow-up that preclude outcome measurement, access to care, health behaviours and literacy, and geography between exposure groups. If death is an outcome, clearly describe the source from which the death record was obtained and how death was verified. If a validated death registry for the study population of interest was not used (e.g., Canadian Vital Statistics), explain why not.

Specify whether a negative control outcome was used. A negative control outcome in this context is an outcome that is not expected to be related to the intervention or exposure.[58] Therefore, the control outcome should not be associated with the intervention or exposure in the study results. An association between the exposure of interest and the control outcome suggests that bias may be responsible for the primary study results. Use of a control outcome can strengthen confidence in study findings. For example, influenza vaccination uptake might be considered as a control outcome in a study of statin use (versus no use) on myocardial infarction risk. If a control outcome is used, justify how it can be considered to be unrelated to the exposure of interest.

Section 7: Summary of Recommendations

- Report definitions for all study outcomes (primary, secondary, and exploratory), where possible.
- Provide a rationale for the outcomes studied and discuss relevant outcomes not included in the study. Consider the use of a COS if 1 is available for the condition of interest under study.
- Provide information about the validity of all outcome definitions.
- Describe whether the timing of the outcome can be accurately measured.
- Specify whether the outcome studied is a surrogate measure of a clinical (patient-centred) outcome and, if so, the strength of the relationship between the surrogate outcome and major clinical outcome(s) of interest.
- Discuss whether outcome misclassification could occur between treatment groups.
- Report whether a negative control outcome(s) was used and justify the negative control outcome(s) selected.

Section 8: Bias, Confounding, and Effect Modifiers or Subgroup Effects

Overview

For effectiveness and safety analyses, bias is systematic error that results in an incorrect estimate of the association between the exposure and the outcome. Unlike RCTs, in RWE studies, the treatment assignment is not controlled and therefore often subject to confounding and other biases related to nonrandom exposure assignment and follow-up of subjects in routine clinical practice. Bias and confounding are critical issues that can hinder the use of RWE for decision-making, and thus studies must have substantial detail provided on potential biases and methods to attempt to address or understand the impact of bias in the study.

A detailed description of each type of bias is outside the scope of this document, but many of these biases are well described elsewhere.[59,60] In brief, the main types of bias in observational studies are information bias, selection bias, and bias from confounding. Information bias arises when key study variables (e.g., exposures or outcomes) are measured differentially between treatment groups. For example, immortal time bias may occur in RWE when follow-up time is included during which the study outcome cannot occur.[61] Selection bias occurs when the inclusion, exclusion, or retention (follow-up) of participants is different between exposure groups. Finally, confounders are factors that are associated with the exposure of interest and the outcome and can therefore induce spurious associations between the treatment of interest and outcome comparing medication and device exposures.[62]

This section requires a critical review and reporting of assumptions. Here, reporting must include differences between the treatment groups' baseline characteristics and the potential for bias. When a risk of bias exists, discuss any methods (i.e., design conduct or analysis) implemented to mitigate or account for bias. Explicitly report if there is a risk of a bias but it was not able to be addressed (e.g., an important missing confounder). Regardless of whether methods were employed to mitigate or account for a bias, investigators must explicitly discuss how they hypothesize that results would be impacted by each identified bias separately (including magnitude and direction of effect toward or away from the null value, if possible).

Effect modification occurs when the measure of association of interest changes over levels of a variable (e.g., an odds ratio that is 20% different between males and females).[32] Given the large study populations and diversity of persons included in studies of RWE, effect modification may be explored to identify heterogeneity of treatment effects and subgroups that may have different risks or benefits from the treatment(s) of interest. At minimum, an exploration of effect modification by main demographic variables (i.e., age, sex, and race or ethnicity) and any other established effect modifiers from the literature should be explored.

Specific Considerations and Recommendations

Bias

Clearly describe any efforts to address potential sources of bias via the study design (e.g., restriction, matching) or statistical analyses. Describe the assumptions or biases that could have influenced the outcomes of the analyses (with direction of the anticipated effect). As applicable, describe the potential for

differential exclusion, exposure measurement, loss to follow-up, and informative censoring, with potential implications. Sensitivity analyses used to test *specific* assumptions and potential biases are described in more detail in section 11.

Confounding

Specify variables that were considered to be potential confounders in the analysis. Specify whether any potential important or relevant confounders could not be measured, the anticipated impact of these confounders on results, and whether data linkage was explored to provide additional information on missing potential confounders. Discuss whether selected confounders were informed by their relationships between the exposure and outcome (e.g., from established literature or clinical expertise). Consider using a causal diagram to illustrate confounders that would be expected to have the strongest relationship between exposures and outcomes.[63] Variables that represent confounders should ideally be measured at or before the exposure to avoid adjusting for intermediaries (factors caused by the exposure that in turn cause the outcome).[64] If these variables are measured after the start of the exposure, they should be clearly indicated as proxies for pre-exposure variables, and their use must be carefully justified (i.e., describing why they could be reasonably not expected to be intermediaries between exposure and outcome). Describe the distribution of potential confounding variables between treatment groups and their comparability at baseline equivalence (e.g., using standardized mean differences). Discuss whether time-varying confounding was considered, especially if participants could switch between the treatment and comparator groups.

Sensitivity Analyses

Sensitivity analyses are defined as the "...analyses conducted with the intent to explore the robustness of inferences from the main estimator to deviations from its underlying modelling assumptions and limitations in the data."[65] Thus, sensitivity analyses (sometimes called stability analyses) test key assumptions and decisions on which results are based. Sensitivity analyses should explore the robustness of effect estimates in relation to deviations in the exposure, outcome definitions, potential unmeasured confounders, and limitations of the data source. It is strongly recommended to test key assumptions on which the primary analysis and results are based using sensitivity analyses. If sensitivity analyses for these assumptions are not feasible or appropriate, explain why.

Quantitative bias analyses (QBAs) are sensitivity analyses that use methods that estimate the direction, magnitude, and uncertainty of study results due to bias.[66,67] QBAs have several benefits, including identifying sources of systemic error and providing ranges of potential impacts of bias on study results, reducing undue confidence in results and conclusions.[53,68] For a more in-depth discussion of the methods and implications of QBAs, refer to established texts.[67,69]

Report all methods used for sensitivity analyses and specify whether each sensitivity analysis was defined a priori or post hoc. Specify the purpose and rationale of each sensitivity analysis, explicitly linking each analysis to a specific assumption or potential bias.

Effect Modification

Specify any known or hypothesized effect modifiers of the effect of the treatment(s) on the outcomes of interest. Describe if any effect modification analyses were conducted, whether these analyses were specified a priori, and what the goals of these analyses were. Importantly, relevant effect modification or subgroup analyses should be identified and conducted based on a prespecified rationale, such as evidence from the literature, previous studies, or a biological rationale. Describe the methods used to examine the subgroups and interactions. Present effect measures for separate subgroups defined by the effect modifiers. If effect modification or subgroup analyses were not used, justify why these analyses were not performed.

Section 8: Summary of Recommendations

- Report all procedures used to address potential sources of bias.
- Specify how potential sources of bias could influence the outcomes of the analyses.
- Specify variables that were considered known or potential confounders in the analysis.
- Describe how confounder variables were selected and if they were informed by a causal diagram.
- Describe and compare the distribution of measured baseline confounding variables between treatment groups.
- Report whether any potential confounders could not be measured and specify the anticipated impact of these confounders on study results.
- Report whether time-varying confounding was considered and, if not, justify why not.
- Specify the methods used to conduct sensitivity analyses that test key assumptions and limitations of the data, and if no sensitivity analyses were conducted, explain why not.
- Specify known or potential effect modifiers.
- Describe any effect modification or subgroup analyses that were conducted and if they were specified a priori. Include if they were identified and conducted based on prespecified rationale such as previous studies or biological rationale. If no effect modification or subgroup analyses were used, justify why they were not needed.
- Report whether effect modification or subgroup analyses were used and describe the methods, and present separate results for each subgroup.

Section 9: Statistical Methods

Overview

The choice of statistical methods and inherent analytic assumptions can greatly impact study findings; however, statistical methods are often underreported in RWE. This guidance does not prescribe or recommend certain statistical methods in general; instead, this section focuses on important reporting principles for methods used, based on the specific research question, data source, and other study specifics. For statistical reporting, transparency in the methods used to generate results is critical. Provision of all or

at least part of the statistical code used is 1 straightforward method to facilitate transparency. If some or all of the statistical code used for a study cannot be provided, justification should be provided as to why not. In addition, this section should provide enough detail on statistical methods used that replication would theoretically be possible without the code. In developing the method(s) of statistical analyses for the primary and secondary outcomes, the estimand principle should be applied as much as possible.[65] The statistical methods used should be decided a priori to ensure that they are not data-driven. Any changes to the statistical methods should be documented before the analyses being conducted, and reasons for the change should be documented. Approaches for handling clinically relevant intercurrent events through application of the estimand framework[70] and missing data should be specified a priori. Providing information and rationale on the sensitivity analyses to test key assumptions and limitations of the study is also important. Finally, precision of effect measures (e.g., confidence intervals) should be provided, as they are more informative than estimates of statistical significance alone for interpreting study results.

Specific Considerations and Recommendations

Essential Statistical Reporting

Indicate any software used for the statistical analyses, including software package, version, and settings, if applicable. Provide the statistical code used for the analysis that allows for replication, if possible. If statistical code cannot be provided, explain why not. Report all statistical methods and models applied to the study and justify each. Report whether a more appropriate, alternate statistical method could have been used and provide rationale as to why it was not conducted (e.g., limited sample size). For example, in studies utilizing regression analysis, report which variables were included and provide detail on how they were operationalized in models (e.g., categorical, continuous, binary). Describe methods used for variable selection (e.g., criteria for stepwise selection of variables in multivariable models, a priori selection of variables). Detail any statistical methods used to control for confounding and to account for missing data, if applicable.

Suitable methods for handling missing data should be fully described and should include the handling of missing data under the primary analysis as well as sensitivity analysis to assess the robustness of the results. Methods for handling missing data should be described and, if applicable, assumptions for the statistical methods for missing data imputation should be checked.[71-73] The appropriateness of the method or methods selected is dependent on many factors, such as the quality of the imputation models and study size. Report the methods that were undertaken to handle missing data and provide rationale for their use. Any assumptions about the data and analyses should be clearly stated and preferably validated when possible (e.g., normal distribution of a variable). In addition, sensitivity analyses can help assess certain statistical assumptions.

Method-Specific Statistical Reporting

Describe methods used to account for differential follow-up time between exposure groups. Report methods used to identify strata and any stratification approaches used. Specify methods used to examine subgroups and interactions. For studies using propensity score methods, report the methods used to construct propensity scores; assumptions underlying the construction of propensity scores or their derivations (e.g.,

inverse probability weights); and details for matching, trimming, weighting, and propensity score diagnostics (e.g., histograms, comparisons of weighted means; and the distribution of propensity scores and/or inverse probability weights), as applicable. Instrumental variables should be used with caution given the strong assumptions required for this method.[74] If an instrumental variable is used, report methods used to assess the validity of the instrument.[74] Report any methods used to combine results of studies or results from different populations, such as using meta-analytical methods.

Statistical Significance and Precision of Estimates

Indicate thresholds of statistical significance. However, investigators should not rely on statistical significance alone for presentation and interpretation of study findings.[75] Instead, estimates of precision should be quantified (e.g., via confidence intervals). Specifically, authors should not describe results as "statistically significant" or "nonsignificant," or rely on thresholds for P values, but rather report the exact P value together with an estimate of precision like a confidence interval. Finally, because the probability of detecting significant effects increases as the number of statistical tests performed increases,[76] report how multiplicity (i.e., multiple testing) was handled, where applicable.

Section 9: Summary of Recommendations

- Indicate the software used for the statistical analysis including software package, version, and analytic tools employed (e.g., macros).
- Provide access to the statistical code used, or if the code cannot be shared, explain why not.
- Report all statistical methods used and justify their selection, including, as applicable:
  - all variables included in regression models
  - the method of variable selection for regression models
  - methods used to control for confounding
  - methods used to account for missing data
  - how follow-up time and changes in exposures were handled
  - subgroup analyses and effect modification
  - stratification, propensity score estimation and assumptions, meta-analysis methods, and validity of instrumental variables.
- Quantify the precision of all estimates using confidence intervals.
- Report the threshold of statistical significance used.

Section 10: Study Findings

Overview

Central to any RWE study is the transparent and accurate reporting of study results. Ultimately, the reported results should align with the study objectives and/or hypotheses described in the methods. Results should

include the estimated effect measures and measures of precision (e.g., 95% confidence intervals) for all primary and secondary outcomes, where applicable. In addition, the numbers of outcome events or summary measures of outcomes (or exposures in case-control studies) are needed. Absolute and relative effect measures, unadjusted and confounder-adjusted estimates, and measures of precision (e.g., 95% confidence intervals) should be clearly reported as applicable. Additionally, these reported values should be accompanied by results from subgroup and sensitivity analyses, and interactions. Analyses that were not planned before starting the study, if conducted, must be clearly presented as post hoc.

Specific Considerations and Recommendations

Reporting Main Analyses

Reporting should be aligned with the prespecified outcomes and methods. Outcomes should be reported in the manner and order in which they were presented in the methods. It is imperative that the outcomes of a study are presented in an objective manner (i.e., without editorializing descriptors like "a major benefit was identified"), providing a comprehensive and accurate description of the findings. Results should be summarized with reference to each study objective and/or hypothesis as described in the methods section of the study. All primary and secondary outcomes (and exploratory outcomes, if applicable) delineated by treatment or exposure groups, their estimated effect measures, and measures of precision (e.g., 95% confidence intervals) should be reported. In addition, numbers of outcome events or summary measures of outcomes (or exposures in case-control studies) must be clearly presented. Confidence intervals are important tools that provide an understanding of the precision of study results and thus should be included, where applicable.

For binary outcomes, absolute and relative effect measures — including measures of precision — are recommended (e.g., median survival estimates and hazard ratios for a time-to-event analysis) to facilitate the interpretability and impact of results, if possible. Report both unadjusted and adjusted estimates, including their measure of precision and the confounders used for adjustment, if applicable.

Reporting Other Analyses

Selective reporting of results is not best practice. All other prespecified analyses that were conducted — such as subgroup analyses, sensitivity analyses, and interactions — should be reported alongside their results, including measures of precision, if applicable. Describe and present any unplanned analyses performed secondarily (not defined a priori), such as subgroup analyses or investigation of alternative exposure categories, and indicate these as post hoc.

Section 10: Summary of Recommendations

- Summarize key results (estimated effect measures, measures of precision) with reference to each study objective and/or hypothesis for primary and secondary outcomes and delineate these results by each treatment or exposure group.
- Provide numbers of outcome events or summary measures of outcomes (or exposures in case-control studies).

- Report both absolute and relative effect measures for binary outcomes, including their measure of precision.
- Report category boundaries when continuous variables are categorized, and consider translating estimates of relative risk into absolute risk.
- Report unadjusted and adjusted estimates, including their measure of precision and confounders used for adjustment.
- Report other prespecified analyses conducted (e.g., subgroup analyses, interactions, sensitivity analyses).
- Describe any unplanned analyses performed secondarily (not defined a priori) and indicate these as exploratory.
- Avoid selective reporting of results.

Section 11: Interpretation and Generalizability

Overview

A thoughtful and balanced interpretation of study results is critical to any RWE study. Primary, secondary, and exploratory study findings should be discussed, including adjusted and unadjusted analyses. Investigators should discuss how the interpretation of results might be affected by the limitations of the study (e.g., bias, confounding, missing data). This section should also include a discussion of the study findings as they relate to similar studies and other relevant evidence. It should also provide a realistic interpretation of the clinical significance of results contextualized within the current literature. Although RWE is often based on a broad range of patients, which can translate into better generalizability, this section should also include considerations of generalizability of study results specific to the Canadian context. Finally, outlining any patient and/or caregiver involvement supporting the interpretation and generalizability of study findings to a Canadian context should be provided.[77,78]

Specific Considerations and Recommendations

Interpretation of Study Results

Interpret the study findings by summarizing the key results from a priori and post hoc primary and secondary analyses. Ensure that causality is not inappropriately inferred from an association. Provide an overall interpretation of results considering the study's objectives, limitations, results from similar studies, and other relevant evidence. Specify implications for clinical practice (clinical significance) in addition to statistical significance. Also summarize key results and an interpretation of unadjusted versus adjusted analysis, if conducted, and discuss the precision of the estimated effect measure(s). Discuss approaches undertaken throughout the study to mitigate potential biases, misclassifications, and/or heterogeneity that could affect study results, to allow for appropriate clinical interpretation of findings. For example, describe the sensitivity of inferences to missing data methods and assumptions. Finally, for studies of adverse events, interpret

results in relation to their impact on the benefit-risk balance of the concerned product(s), the clinical context of the safety issue, and the risk management plan of the product(s), if applicable.

Generalizability

Discuss the generalizability (external validity) of the study results, considering the data source, characteristics of the final study population versus the population in Canada, considerations of equity and diversity of participants,[43] and inclusion and exclusion criteria. Investigators should acknowledge whether some participant groups are not well represented — or conversely, are over-represented — and if either is the case, how under- or over-inclusion of these groups may impact the generalizability of the study findings. Additionally, it is suggested that investigators consider potential variation in quality of care and access to the intervention when discussing generalizability, as these considerations may affect external validity of the study results. Discuss the study findings in relation to differences in the treatment pathways or care settings seen in the analytical sample and the Canadian health care system, as it may impact on the relevance of results to the Canadian context.

Section 11: Summary of Recommendations

- Provide an interpretation of the primary and secondary study results, as applicable.
- Interpret the findings from adjusted and unadjusted results, as applicable.
- Discuss the precision of the effect measure(s).
- Discuss how potential biases and sensitivity of study assumptions may impact the results and subsequent interpretation.
- Discuss the implication(s) of findings for clinical practice, including the risk-benefit profile of the treatment, if applicable.
- Interpret study findings in relation to current literature.
- Discuss the generalizability (external validity) of study results to the population in Canada.

Section 12: Limitations

Overview

All studies have limitations and should be discussed. Much of this section refers to limitations previously mentioned in this document (e.g., data limitations acknowledged in section 3). This section should also include considerations of limitations of the data, sample size, generalizability, and clinical significance of results, in addition to typical discussions of bias and confounding. Limitations mentioned should be comprehensive; for each limitation, include a discussion of how the limitation may change study results or interpretation.

Specific Considerations and Recommendations

Data Limitations

Discuss the implications of using data that were not created or collected to answer the specific study question(s). For example, describe the degree to which the chosen databases adequately capture the drug exposure of interest. Discuss any limitations arising from study variables that were constructed by combining multiple data elements (including both structured and unstructured data), or come from different linked data sources (e.g., response rates, missing or incomplete data, and necessary imputations applied).

Bias and Confounding

Discuss sources of potential bias or imprecision, including their direction and magnitude of effect on study results. Discuss any misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as applicable. Discuss the potential for confounding by indication, contraindication, or disease severity; selection bias; or other forms of bias reported as part of section 8 as alternative explanations for the study findings. Explicitly report whether the results are plausible, given the observed magnitude of effect, design and data limitations, opportunities for the influence of bias, chance, or confounding. Also consider including the plausibility of results by using causality frameworks, such as the Bradford Hill criteria[79] (e.g., timing, dose-response, biological plausibility, consistency).

Other Limitations

Discuss the precision of study findings and whether imprecision is a limitation of results. Report whether the observed results are clinically relevant, regardless of whether they are statistically significant. Statistical significance alone does not exclusively determine the clinical importance of the findings because some registries include large amounts of health care data, and very small effect measures can be statistically significant without having meaningful implications for clinical practice.

Section 12: Summary of Recommendations

- Provide a consideration of limitations of the study, including the data source, missing data, bias and confounding, imprecision or sample size limitations, and whether results are clinically meaningful.
- Discuss the plausibility of results and whether results could be due solely to bias, chance, or confounding.

Forward-Looking Statement and Conclusion

CADTH has partnered with Health Canada, INESSS, and other health system stakeholders to advance the integration of RWE into decision-making. *Guidance for Reporting Real-World Evidence* forms the foundation for transparent reporting of RWE studies and facilitates appraisal of RWE for the purpose of supporting decision-making. It outlines principles that are consistent with regulatory and HTA standards, both in Canada and internationally.

This guidance document was developed during a time of immense change in the fields of RWD and RWE, and the expectation is that this document will be periodically updated or expanded over time as this area evolves. *Guidance for Reporting Real-World Evidence* allows flexibility to accommodate both the heterogeneous nature of RWE and its rapid evolution, while ensuring that studies are sufficiently detailed and transparent to facilitate regulatory and HTA appraisal and, ultimately, decision-making.

References

1. Ahead of the curve: shaping future-ready health systems. Ottawa (ON): CADTH; 2022: https://strategicplan.cadth.ca/wp-content/uploads/2022/03/cadth_2022_2025_strategic_plan.pdf. Accessed 2023 Mar 28.
2. Elements of real world data/evidence quality throughout the prescription drug product life cycle. Ottawa (ON): Government of Canada; 2020: https://www.canada.ca/en/services/health/publications/drugs-health-products/real-world-data-evidence-drug-lifecycle-report.html. Accessed 2023 Mar 28.
3. CADTH. Real-world evidence: a primer. 2022; https://www.cadth.ca/real-world-evidence-primer, 2022 Oct 7.
4. Government of Canada. Optimizing the use of real world evidence to inform regulatory decision-making: health products and food branch notice. 2019; https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/announcements/optimizing-real-world-evidence-regulatory-decisions.html, 2022 Oct 7.
5. Plamondon G. State of knowledge: integration of real-world data and evidence to support decision-making in the pharmaceutical sector. Québec (QC): INESSS; 2022: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Medicaments/INESSS_Real_world_data_SK.pdf. Accessed 2022 Oct 7.
6. Sobel R, Girman C, Ehrenstein V, Nyberg F, Soriano-Gabarró M, Toh D. ISPE's position on real-world evidence (RWE). Bethesda (MD): International Society for Pharmacoepidemiology; 2020: https://pharmacoepi.org/pub/?id=136DECF1-C559-BA4F-92C4-CF6E3ED16BB6. Accessed 2022 Oct 7.
7. Capkun G, Corry S, Dowling O, et al. Can we use existing guidance to support the development of robust real-world evidence for health technology assessment/payer decision-making? *Int J Technol Assess Health Care.* 2022;38(1):e79. PubMed
8. Ali MS, Prieto-Alhambra D, Lopes LC, et al. Propensity Score Methods in Health Technology Assessment: Principles, Extended Applications, and Recent Advances. *Front Pharmacol.* 2019;10:973. PubMed
9. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med.* 2000;342(25):1887-1892. PubMed
10. Yuan H, Ali MS, Brouwer ES, et al. Real-World Evidence: What It Is and What It Can Tell Us According to the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG). *Clin Pharmacol Ther.* 2018;104(2):239-241. PubMed
11. Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf.* 2017;26(9):1033-1039. PubMed
12. Murphy G, de Léséleuc L, Kaunelis D, Adcock L. Use of real-world evidence in single-drug assessments. *(Environmental scan no. 74)*. Ottawa (ON): CADTH; 2018: https://www.cadth.ca/sites/default/files/pdf/es0323-rwe-in-single-drug-appraisal.pdf. Accessed 2022 Oct 7.
13. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ.* 2015;350:h2147. PubMed
14. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol.* 2009;62(5):464-475. PubMed
15. Hernán MA, Robins JM. *Causal Inference. What If.* Boca Raton (FL): Chapman & Hall/CRC; 2020.
16. Guyatt G RD, Meade MO, Cook DJ. eds. *Users' guides to the medical literature: a manual for evidence-based clinical practice.* New York (NY): McGraw Hill; 2015: https://jamaevidence.mhmedical.com/content.aspx?bookid=847&sectionid=6903071https://jamaevidence.mhmedical.com/content.aspx?bookid=847&sectionid=69030714. Accessed 2022 Aug 30.
17. Consort. Transparent reporting of trials. 2010; http://www.consort-statement.org/. Accessed 2022 Aug 31.
18. Equator Network. Enhancing the quality and transparency of health research. [no date][no date]; . Accessed 2022 Aug 31.
19. Prisma. Transparent reporting of systematic reviews and meta-analyses. 2020; https://www.prisma-statement.org/. Accessed 2022 Aug 31.

20. STROBE. Strengthening the reporting of observational studies in epidemiology. 2022; https://www.strobe-statement.org/. Accessed 2022 Aug 31.
21. RECORD. REporting of studies Conducted using Observational Routinely-collected data: Record-PE checklist. 2019; https://www.record-statement.org/checklist-pe.php. Accessed 2022 Aug 31.
22. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol.* 2016;183(8):758-764. PubMed
23. Hernan MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA.* 2022;328(24):2446-2447. PubMed
24. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol.* 2016;79:70-75. PubMed
25. Wang SV, Pinheiro S, Hua W, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ.* 2021;372:m4856. PubMed
26. Wang SV, Pottegård A, Crown W, et al. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force. *Pharmacoepidemiol Drug Saf.* 2023;32(1):44-55. PubMed
27. Center for Open Science. Real world evidence registry. 2023; https://osf.io/registries/rwe/discover. Accessed 2023 Apr 3.
28. U.S. National Library of Medicine. ClinicalTrials.gov. 2023; https://clinicaltrials.gov/. Accessed 2023 Apr 3.
29. The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). Amsterdam (NL): European Medicines Agency; 2023: https://www.encepp.eu/encepp/studiesDatabase.jsp. Accessed 2023 Apr 3.
30. Canadian Institutes of Health Research. Strategy for Patient-Oriented Research - Patient Engagement Framework. https://cihr-irsc.gc.ca/e/48413.html. Accessed 2023 May 4.
31. Oehrlein E, Schoch S, Burcu M, et al. Developing Patient-Centered Real-World Evidence: Emerging Methods Recommendations From a Consensus Process. *Value in Health.* 2023;26. PubMed
32. International Committee of Medical Journal Editors. Defining the Role of Authors and Contributors. 2023; https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html. Accessed 2023 Apr 20.
33. Jaksa A, Arena PJ, Chan KKW, Ben-Joseph RH, Jónsson P, Campbell UB. Transferability of real-world data across borders for regulatory and health technology assessment decision-making. *Front Med.* 2022;9. PubMed
34. Kahn MG, Callahan TJ, Barnard J, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS (Wash DC).* 2016;4(1):1244. PubMed
35. Daniel G, Silcox C, Bryan J, McClellan M, Romine M, Frank K. Characterizing RWD quality and relevancy for regulatory purposes. Washington (DC): Duke-Margolis Center for Health Policy; 2018: https://healthpolicy.duke.edu/sites/default/files/2020-03/characterizing_rwd.pdf. Accessed 2022 Aug 20.
36. The Observational Medical Outcomes Partnership Common Data Model Working Group. Observational Medical Outcomes Partnership (OMOP) Common Data Model. 2023; https://ohdsi.github.io/CommonDataModel/. Accessed 2023 Mar 28.
37. CIHI Reference Data Model Toolkit. Ottawa (ON): Canadian Institute for Health Information; 2022: https://www.cihi.ca/sites/default/files/document/cihi-reference-data-model-toolkit-en.pdf. Accessed 2023 Apr 3.
38. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 10). EMA/95098/2010.
39. Sayers A, Ben-Shlomo Y, Blom AW, Steele F. Probabilistic record linkage. *Int J Epidemiol.* 2016;45(3):954-964. PubMed
40. Guilhaume C. A tool to assess the registries quality: The Registry Evaluation and Quality Standards Tool (REQueST). *Eur J Public Health.* 2021;31(Supplement_3).
41. Center for Open Science. How to make a data dictionary. 2022; https://help.osf.io/article/217-how-to-make-a-data-dictionary. Accessed 2023 Mar 28.

42. Schulz KF, Altman DG, Moher D, Consort Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010;152(11):726-732. PubMed

43. Coe I, Gaensler B, Ghose S, Kerr J, Ronksy J, Smith M. Guide for applicants: considering equity, diversity and inclusion in your application. Ottawa (ON): Natural Sciences and Engineering Research Council of Canada; 2017: https://www.nserc-crsng.gc.ca/_doc/EDI/Guide_for_Applicants_EN.pdf. Accessed 2022 Oct 7.

44. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol.* 2003;158(9):915-920. PubMed

45. Canadian Institutes of Health Research. Key considerations for the appropriate integration of sex and gender in research. 2019; https://cihr-irsc.gc.ca/e/50835.html. Accessed 2023 Mar 28.

46. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28(25):3083-3107. PubMed

47. Mamdani M, Sykora K, Li P, et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. *BMJ.* 2005;330(7497):960-962. PubMed

48. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. *Am Stat.* 2016;70(2):129-133.

49. Mustafina RH. Enhancement of the radiation effect by fibrinolysin and quantitative pattern of the change of tumor radiosensitivity. *Neoplasma.* 1984;31(3):263-269. PubMed

50. Carrigan G, Bradbury BD, Brookhart MA, et al. External Comparator Groups Derived from Real-world Data Used in Support of Regulatory Decision Making: Use Cases and Challenges. *Curr Epidemiol Rep.* 2022;9(4):326-337.

51. Gray CM, Grimson F, Layton D, Pocock S, Kim J. A Framework for Methodological Choice and Evidence Assessment for Studies Using External Comparators from Real-World Data. *Drug Saf.* 2020;43(7):623-633. PubMed

52. Han B, Zhan J, John Zhong Z, Liu D, Lindborg S. Covariate-adjusted borrowing of historical control data in randomized clinical trials. *Pharm Stat.* 2017;16(4):296-308. PubMed

53. National Institute for Health and Care Excellence. NICE real-world evidence framework. 2022: https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837. Accessed 2023 Apr 20.

54. Mercon K, Lallingeer K, Mahendraratnam N, et al. A roadmap for developing study endpoints in real world settings. Washington (DC): Duke-Margolis Center for Health Policy; 2020: https://healthpolicy.duke.edu/sites/default/files/2020-08/Real-World%20Endpoints.pdf. Accessed 2022 Oct 27.

55. Velentgas P DN, Nourjah P, Smith SR, Torchia MM, eds. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville (MD): Agency for Healthcare Research and Quality; 2013: https://www.ncbi.nlm.nih.gov/books/NBK126190/. Accessed 2023 Apr 3.

56. Johnston BC PD, Devji T, et al. Chapter 18: Patient-reported outcomes. In: Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA ed. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*. London (GB): Cochrane; 2022: https://training.cochrane.org/handbook. Accessed 2022 Oct 27.

57. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials.* 2012;13:132. PubMed

58. Arnold BF, Ercumen A. Negative Control Outcomes: A Tool to Detect Bias in Randomized Trials. *JAMA.* 2016;316(24):2597-2598. PubMed

59. Rothman K, Greenland S, Lash TL. *Modern Epidemiology.* 3rd ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2008.

60. Strom BL, Kimmel SE, Hennessey S. *Pharmacoepidemiology.* Sixth ed. New York (NY): John Wiley & Sons Ltd; 2019.

61. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol.* 2008;167(4):492-499. PubMed

62. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. *JAMA.* 2016;316(17):1818-1819. PubMed

63. VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol.* 2007;166(9):1096-1104. PubMed

64. Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clin Epidemiol.* 2017;9:331-338. PubMed

65. ICH E9(R1) Expert Working Group. ICH E9(R1) estimands and sensitivity analysis in clinical trials. Geneva (CH): International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2021: https://database.ich.org/sites/default/files/E9%28R1%29%20Training%20Material%20-%20PDF_0.pdf. Accessed 2022 Oct 25.

66. Gustafson P. *Measurement Error and Misclassification in Statistics and Epidemiology Impacts and Bayesian Adjustments.* Boca Raton (FL): Chapman & Hall/CRC; 2004.

67. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data.* New York (NY): Springer; 2009.

68. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol.* 2014;43(6):1969-1985. PubMed

69. Lash TL, Fox MP, Cooney D, Lu Y, Forshee RA. Quantitative Bias Analysis in Regulatory Settings. *Am J Public Health.* 2016;106(7):1227-1230. PubMed

70. Luijken K, van Eekelen R, Gardarsdottir H, Groenwold RHH, van Geloven N. Tell me what you want, what you really really want: Estimands in observational pharmacoepidemiologic comparative effectiveness and safety studies. *Pharmacoepidemiol Drug Saf.* 2023(epub ahead of print). PubMed

71. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res.* 2007;16(3):219-242. PubMed

72. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med.* 2011;30(4):377-399. PubMed

73. Carpenter JR, Kenward MG, White IR. Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Stat Methods Med Res.* 2007;16(3):259-275. PubMed

74. Davies NM, Smith GD, Windmeijer F, Martin RM. Issues in the reporting and conduct of instrumental variable studies: a systematic review. *Epidemiology.* 2013;24(3):363-369. PubMed

75. Halsey LG. The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biol Lett.* 2019;15(5):20190174. PubMed

76. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: The perils of multiple testing. *Perspect Clin Res.* 2016;7(2):106-107. PubMed

77. CADTH. CADTH framework for patient engagement in health technology assessment. 2022; https://www.cadth.ca/cadth-framework-patient-engagement-health-technology-assessment. Accessed 2022 Oct 7.

78. Canadian Institutes of Health Research. Canada's Strategy for Patient-Oriented Research: improving health outcomes through evidence-informed care. 2012; https://cihr-irsc.gc.ca/e/44000.html. Accessed 2023 Apr 25.

79. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med.* 1965;58(5):295-300. PubMed

Appendix 1: Methods

Methods

A 3-phase process to develop this guidance document was used. Phase 1 aimed to leverage and extend 2 existing environmental scans of international RWE evidence[5,12] and an evidence-mapping process to develop candidate items to be included in the guidance. Phase 2 used a modified Delphi process with an Expert Methods Panel to select final recommendations from the list of candidate recommendations, include additional relevant items if required, and include considerations for operationalizing the recommendations as well as special considerations for the Canadian context, as appropriate. Phase 3 implemented a stakeholder consultation plan and Expert Methods Panel survey to incorporate and finalize revisions to the guidance based on public and stakeholder feedback.

Phase 1, Part 1: Identification of Documents on RWE

Potential articles related to Canadian and international agency guidance, reporting tools, and policy statements on RWD and RWE were first identified through review of 2 existing environmental scans: a state of knowledge report by the Institut national d'excellence en santé et en services sociaux (INESSS)[5] and a 2020 Environmental Scan by CADTH, *Use of Real-World Evidence in Single-Drug Assessments*.[12] Detailed methods used in these environmental scans are described in the original documents. Briefly, in the Environmental Scan conducted by CADTH, authors conducted a literature search to identify relevant guidelines or policy papers from government agencies through searching of standard databases (OVID MEDLINE, PubMed) and HTA or regulatory agency websites. Then, a supplemental survey was sent to a subset of agencies hosting drug review program to identify additional documents. Similarly, the report published by INESSS conducted a literature search of standard databases and sources of grey literature to identify relevant documents. Then, a review of the references and keywords of identified documents was conducted to identify additional candidate documents.

Next, these environmental scans were extended to identify potential additional documents by using a citation-search method and consulting the Expert Methods Panel to identify articles that had not yet been included or were currently in development. In total, 37 documents were identified for review and data extraction (Appendix 2).

Phase 1, Part 2: Extracting Candidate Recommendations for RWD and RWE From Identified Documents (Evidence Mapping)

Two data extraction tools (matrices) were created to organize identified recommendations. Recommendations were categorized based on RWE reporting and RWE conduct (methods considerations). For each category, we developed a matrix with subcategories (e.g., protocol, exposures), to which recommendations were mapped or organized. Additional matrix subcategories were added if a new theme was identified. Two investigators (KH, TA) independently reviewed all identified documents and

independently extracted data on recommendations on the reporting and conduct of RWE. A third investigator (MT) reviewed the extracted data for accuracy.

Table 1: Categories of Data Extraction Matrices

Component	Categories to which recommendations were mapped
Reporting of real-world evidence	Study design, setting, participants, study size, variables and definitions, data sources/ management, reporting on follow-up time, data access and cleaning methods, data linkage, bias, statistical methods, adverse event reporting, deviations from protocol, data transformations, governance, statistical software, participant consent, minimum dataset requirements, quality assurance, data security, data codes, reporting on participants, descriptive data, outcome data, main results, other analyses, limitations, interpretation, generalizability, reliability, presentation of results, financing
Conduct of real-world evidence	Data quality, data appropriateness/quality/fitness of use, generalizability, data cleaning/ dataset creation, study team, protocols/registry/study planning, publication bias, study question/objective/appropriateness, study design, study population, exposure/exposure definitions, controls/comparators, outcomes, exposure-outcome risk window and follow-up, causality/confounders/bias/sensitivity analyses, effect modifiers and subgroup effects, missing data, analysis, interpretation and dissemination of results, other notes, other documents cited

One investigator (MT) removed duplicate recommendations and revised all recommendations with common language (e.g., use of "exposure" versus "drug" or "intervention") to allow mapping of major themes and enhance clarity for the Expert Methods Panel and stakeholders. The other 2 investigators (KH, TA) then independently reviewed the duplicate removal and standardization of language to ensure that items were indeed duplicates and provide consensus on standardized language. All 3 investigators then mapped recommendations to major themes that were revised from the categories in the data extraction matrices. In total, 200 candidate recommendations were extracted, distilled, and mapped to 16 major themes (Table 2).

Phase 2, Part 1: Establishment of the Expert Methods Panel

Authorship and Leadership Teams

The Expert Methods Panel was purposefully selected to include 10 members based in Canada and 5 international members. Experts were selected based on established expertise in the field, evidenced by a publication record of applying RWE and/or developing methods. Selection purposefully aimed for diversity in expertise, geographic location, use of differing data (i.e., administrative, registry), methods expertise (with specific interest in epidemiology and economics), gender, and career level. International experts were also selected based on experience supporting or leading international guidance. All experts had to declare potential conflicts of interests and align with CADTH's conflict of interest policy.

The Expert Methods Panel also included representatives of key Canadian stakeholder organizations (Health Canada, CADTH, INESSS, CIHI, Statistics Canada) who had established expertise in RWE. While these panel members did not complete the survey, they participated in all Expert Methods Panel meetings as part of the consensus process. Two CADTH representatives — the Vice-President of Scientific Advice, Methodologies,

and Resources and co-chair of the Guidance for Reporting Real-World Evidence WG — also attended the meetings as observers.

Phase 2, Part 2: Delphi/Consensus Process

Methods Authorship Team

The Delphi process was led by 3 authors (TA, KH, MT) with experience in pharmacoepidemiology, systematic reviews, and knowledge synthesis. The authors led the data collection and synthesis for Phase 1, attended the consensus meetings as observers (they did not vote in the surveys), and iteratively drafted the resulting guidance document. The consensus process and group discussions were facilitated by a team member (CF) with expertise in Delphi methodology and knowledge translation.

Structure of Modified Delphi and Data Collection

The 200 recommendations were grouped into 16 themes and programmed into an online questionnaire using SurveyMonkey. Canadian and international RWE experts were asked to determine the importance of including each item into the guidance document. Each item was ranked on an anchored scale of 1 to 4 where 1 meant “not important” and 4 meant “very important” for inclusion. Participants had the opportunity to include feedback on each item via an open-ended text box. The survey was circulated to participants by email and they were given 10 days to independently complete it. Two email reminders were sent at 5-day intervals. Prior to the meeting, participants received a list of all 200 items, their scores (with items for discussion flagged), and open-ended comments. Items with a score of 1 or 2 were grouped as “exclude” and those with a 3 or 4 were grouped as “include”. Items that generated 70% or greater agreement to include from respondents were included in the guidance document. The same level of agreement was used for items to exclude. Items that generated less than 70% agreement were discussed in a virtual meeting that took place on June 22, 2022.

During the meeting, participants took part in a facilitated discussion (guided by CF), in which each item that did not generate consensus was discussed. Participants voted on whether to “include,” “omit,” or “revise” each item using an online polling feature. Items that generated 70% or greater consensus were included, omitted, or revised, as per group consensus. Six items that did not generate consensus were put forward for additional asynchronous discussion, via email. Participants voted on each of the items (to include or exclude) and were requested to provide comments to support their decisions within 14 days. Additionally, participants took part in a general discussion about the scope, content, and style of the guidance document. All facilitated discussions were recorded, transcriptions were generated, and the authorship team took detailed notes.

Following the first discussion, the Methods Authorship Team drafted the first iteration of the guidance document. Items confirmed by the Expert Methods Panel were inputted into a reporting checklist and elaborated upon to provide additional guidance on how to implement them. The guidance document was circulated to the Expert Methods Panel for review; the initial review was 2 weeks. Participants were invited to provide feedback on the document via email or in the shared document. Additional feedback was collected

from CADTH and Health Canada. The authorship team (TA, KH, MT) compiled the feedback and a second facilitated discussion (guided by CF) to determine which feedback should be incorporated, and which items should be revised, was held virtually on September 20, 2022. Participants were guided in a facilitated discussion; an online polling feature to "include" or "not include" an item was available as required.

Additionally, participants had in-depth discussions to define scope, content, and style of this document. Further in-depth feedback on these items was collected via email. Discussion points and asynchronous feedback were also collected via email and incorporated into the document as appropriate, resulting in an updated draft of the guidance document.

Results

A total of 13 respondents completed the survey. (Refer to Table 2 for results.) Themes that generated 100% agreement included: *participants, exposure definitions and comparators, effect modifiers,* and *study findings*. Other themes generated reasonable to high levels of agreement (67% to 92%). Themes with low levels of agreement were *variables* and *data access and cleaning methods*.

A total of 29 individuals attended the first discussion meeting (14 voting members, 15 stakeholder representatives or observers). Thirty items were discussed during the June 22, 2022, meeting and 6 items were discussed asynchronously after the call due to time constraints.

A total of 14 participants from the Expert Methods Panel and members of the Leadership Review Team from CADTH, Health Canada, and INESSS reviewed the first iteration of the *Guidance for Reporting Real-World Evidence* document. Fifteen discussion items were put forward during the September 20, 2022, meeting, which was attended by 21 participants (10 voting members, 11 observers). All items generated consensus. Two clarifications regarding wording of recommendations and language of the aims were further discussed asynchronously via email in support of the lead writers of the document.

Table 2: Results From First Methods Expert Panel Consensus Survey

Section (n = 13 of 15 panel members)	Number of questions	Overall agreement, n (%)	Drop
1. Study design and question	22	18 (82)	2
2. Setting and context	11	9 (82)	0
3. Data access and cleaning methods	14	8 (57)	1
4. Data linkage	8	6 (75)	2
5. Data sources/measurement	12	8 (67)	0
6. Participants	22	22 (100)	0
7. Exposure definitions and comparators	12	12 (100)	0
8. Outcomes	18	12 (67)	2
9. Variables (covariates and all variable measurement)	9	4 (44)	0
10. Effect modifiers	3	3 (100)	0

Section (n = 13 of 15 panel members)	Number of questions	Overall agreement, n (%)	Drop
11. Bias and confounding	8	7 (88)	2
12. Statistical analysis	19	15 (79)	0
13. Participant characteristics	9	8 (89)	0
14. Study findings	12	12 (100)	0
15. Limitations	9	8 (89)	0
16. Interpretation and generalizability	12	11 (92)	2
Overall	200	163 (82)	11

Phase 3: Stakeholder Consultation and Guidance Revisions

A stakeholder consultation process and public feedback period was implemented to engage with members of the Canadian health technology ecosystem. A draft of the guidance was posted on the CADTH website for public and stakeholder review and feedback for 8 weeks. During this stakeholder feedback period, members of the Methods Authorship Team and Leadership Review Team participated in multiple in-person and virtual events and leveraged established networks to increase visibility of the posted draft report for comment and to offer opportunities to provide additional feedback.

Fifty-four sets of feedback, across various types of stakeholders, were received in response to the consultation and public comment period. Pharmaceutical companies, patient groups, and academic institutions comprised more than 50% of written feedback submissions. The Methods Authorship Team and Leadership Review Team collaboratively reviewed the feedback and grouped comments into general themes that could be applied throughout the document (e.g., consistency of language) or RWE reporting–specific feedback. Major revisions to the document to address RWE reporting–specific feedback were grouped by theme and presented to voting members of the Expert Methods Panel in a survey similar to that used in Phase 2, wherein they voted to include or exclude a revision and could provide additional comments. Major revisions were defined as any revisions that substantially altered the meaning or content of the text. Changes such as rewording a sentence for clarity, adding examples, adding extra citations, changing a figure or table caption, or adjusting the formatting of the manuscript were considered minor. An item was included or excluded if there was 70% or greater consensus from the Expert Methods Panel on the survey item. Items not reaching consensus or that were flagged were discussed in a meeting with the Expert Methods Panel in a similar fashion to the meetings in Phase 2, and modified as needed to finalize revision of the guidance document. In total, 51 major revisions were included in the survey, and 11 of the 14 Expert Methods Panel members completed the survey (note: 1 expert was ineligible because he switched roles). Major revisions to the guidance are listed in the response document.

All but 8 items (e.g., inclusion of lab or in vitro evidence, machine learning methods, public platform for protocol registration) had expert consensus through the survey on whether to include or exclude. These 8 items were discussed at a virtual meeting with the Expert Methods Panel on March 1, 2023. During the meeting, participants took part in a facilitated discussion (guided by CF) in which each item that did not

generate consensus was discussed. Items that generated 70% or greater consensus were included in or excluded from the guidance document as per Expert Methods Panel consensus. Through the survey and consequent discussions, all items achieved consensus. The Expert Methods Panel agreed to the addition of a new section of the guidance concerning implementation to communicate how the guidance may be leveraged in practice. All facilitated discussions were recorded, transcriptions were generated, and the Methods Authorship Team took detailed notes.

Appendix 2: Documents Reviewed for Candidate Recommendations on RWE or RWD Reporting for Expert Survey

Framework for FDA's real-world evidence program. Silver Spring (MD): U.S. Food and Drug Administration; 2018: https://www.fda.gov/media/120060/download. Accessed 2023 Apr 20.*

Guidance for industry and FDA staff: best practices for conducting and reporting pharmacoepidemiologic safety studies using electronic healthcare data. Silver Spring (MD): U.S. Food and Drug Administration; 2013: https://www.fda.gov/media/79922/download. Accessed 2023 Apr 20.*

Plamondon G. State of knowledge: integration of real-world data and evidence to support decision-making in the pharmaceutical sector. Québec (QC): INESSS; 2022: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Medicaments/INESSS_Real_world_data_SK.pdf. Accessed 2022 Oct 7.*

Guidance for registry-based studies. Amsterdam (NL): European Medicines Agency; 2021: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en-0.pdf. Accessed 2023 Apr 20. *

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) guide on methodological standards in pharmacoepidemiology. Amsterdam (NL): European Medicines Agency; 2021: https://www.encepp.eu/standards_and_guidances/documents/1.ENCePPMethodsGuideRev.9.pdf. Accessed 2023 Apr 20.*

Guidance for the format and content of the final study report of non-interventional post authorisation safety studies. Amsterdam (NL): European Medicines Agency; 2013: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-format-content-final-study-report-non-interventional-post-authorisation-safety-studies_en.pdf. Accessed 2023 Apr 20. *

Registries for evaluating patient outcomes: a user's guide. Rockville (MD): Agency for Healthcare Research and Quality; 2020: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-evaluating-patient-outcomes-4th-edition.pdf. Accessed 2023 Apr 20.*

The Registry Evaluation and Quality Standards Tool (REQueST). Diemen (NL): European Network for Health Technology Assessment; 2019: https://www.eunethta.eu/request-tool-and-its-vision-paper/. Accessed 2023 Apr 20.*

Schulz KF, Altman DG, Moher D, Consort Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010;152(11):726-732.* PubMed

von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007;147(8):573-7.* PubMed

Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885.* PubMed

Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532.* PubMed

Orsini, LS, Berger ML, Crown W, et al. Improving transparency to build trust in real-world secondary data studies for hypothesis testing—why, what, and how: recommendations and a road map from the Real-World Evidence Transparency Initiative. *Value Health*. 2020; 23(9):1128-1136. PubMed

Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I, Gagne JJ, Gini R, Klungel O, Mullins CD, Nguyen MD, Rassen JA, Smeeth L, Sturkenboom M; joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032.* PubMed

Oortwijn, W., Sampietro-Colom, L., and Trowman, R. (2019). How to Deal with the Inevitable: Generating Real-World Data and Using Real-World Evidence for HTA Purposes – From Theory to Action. *Int J Technol Assess Health Care*. 35(4), 346-350. PubMed

National Institute for Health and Care Excellence. NICE real-world evidence framework. 2022: https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837. Accessed 2023 Apr 20.*

Health Canada. Elements of real world data/evidence quality throughout the prescription drug product life cycle. 2019: https://www.canada.ca/en/services/health/publications/drugs-health-products/real-world-data-evidence-drug-lifecycle-report.html. Accessed 2023 Apr 20.*

Comparators and comparisons: criteria for the choice of the most appropriate comparator(s), summary of current policies and best practice recommendations. Diemen (NL): European Network for Health Technology Assessment; 2015: https://www.eunethta.eu/wp-content/uploads/2018/03/Criteria_WP7-SG3-GL-choice_of_comparator_amend2015.pdf. Accessed 2023 Apr 20.*

The Use of real-world evidence for medical device assessment – an environmental scan *(Environmental scan no. 91).* Ottawa (ON): CADTH; 2020: https://www.cadth.ca/sites/default/files/es/es0344-use-of-real-world-evidence-for-med-dev-assessment.pdf. Accessed 2023 Apr 20.*

Coles B, Tyrer F, Hussein H, Dhalwani N, Khunti K. Development, content validation, and reliability of the Assessment of Real-World Observational Studies (ArRoWS) critical appraisal tool. *Ann Epidemiol*. 2021;55:57-63.e15.* PubMed

CASP qualitative checklist. Oxford (UK): CASP UK; 2018: https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf. Accessed 2023 Apr 20*

CASP Case-Control Study Checklist. Oxford (UK): CASP UK; 2018: https://casp-uk.net/images/checklist/documents/CASP-Case-Control-Study-Checklist/CASP-Case-Control-Study-Checklist-2018-fillable-form.pdf. Accessed 2023 Apr 20.*

The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa (ON): Ottawa Hospital Research Institute; [no date][no date]: . Accessed 2023 Apr 20.*

Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg*. 2003;73(9):712-6.* PubMed

Liaw S-T, Guo JGN, Ansari S, et al. Quality assessment of real-world data repositories across the data life cycle: A literature review. *J Am Med Inform Assoc.* 2021;28(7):1591-99. PubMed

Kahn MG, Callahan TJ, Barnard J, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS (Wash DC)*. 2016;4(1):1244.* PubMed

Determining real-world data's fitness for use and the role of reliability. Washington (DC): Duke-Margolis Center for Health Policy; 2019: https://healthpolicy.duke.edu/sites/default/files/2019-11/rwd_reliability.pdf. Accessed 2023 Apr 20. *

Daniel G, Silcox C, Bryan J, McClellan M, Romine M, Frank K. Characterizing RWD Quality and Relevancy for Regulatory Purposes. Washington (DC): Duke-Margolis Center for Health Policy; 2018: https://healthpolicy.duke.edu/sites/default/files/2020-03/characterizing_rwd.pdf. Accessed 2022 Aug 20.*

Miksad RA, Abernethy AP. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. *Clin Pharmacol Ther*. 2018;103(2):202-205.* PubMed

Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf.* 2017;26(9):1033-1039.* PubMed

International Society for Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practices (GPP). 2023: https://www.pharmacoepi.org/resources/policies/guidelines-08027/. Accessed 2023 Apr 20.*

Patient-Centered Outcomes Research Institute. Methodology standards checklist. 2023: https://www.pcori.org/document/methodology-standards-checklist. Accessed 2023 Apr 20.*

Dreyer NA, Bryant A, Velentgas P. The GRACE Checklist: A Validated Assessment Tool for High Quality Observational Studies of Comparative Effectiveness. *J Manag Care Spec Pharm*. 2016;22(10):1107-13* PubMed

Lockwood C, Sfetcu R, Oh EG. Synthesizing Quantitative Evidence. Philadelphia (PA): Lippincott-Joanna Briggs Institute; 2011: https://nursing.lsuhsc.edu/JBI/docs/JBIBooks/Syn_Quant_Evidence.pdf. Accessed 2023 Apr 20.*

Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919. * PubMed

Wang SV, Pinheiro S, Hua W, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ*. 2021;372:m4856.* PubMed

Bolislis WR, Fay M, Kuhler TC. Use of Real-world Data for New Drug Applications and Line Extensions. *Clin Therap.* 2020;42(5):926-938. PubMed

** Denotes that candidate recommendations were extracted from the document.*

Appendix 3: Recommendation Checklist

Table 3: Recommendation Checklist

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
Section 1: Study design and research questions	1. Report a clearly stated aim and study question.		
	2. Report the overall study design.		
	3. Provide a rationale for the choice of study design.		
	4. Provide a relevant review of the literature to evaluate pertinent information and gaps in knowledge.		
	5. Describe key elements of the study design (e.g., matching).		
	6. Consider the use of study diagrams to illustrate key aspects of the study design.		
	7. Strongly recommend to develop and reference an a priori protocol.		
	8. Describe all study team members, including the role of patient partners, and any conflicts of interest.		
	9. Describe the study governance structure, especially who was responsible for final decision-making.		
	10. Report any research ethics approval (or equivalent).		
	11. Disclose sources of funding.		
Section 2: Setting and context	1. Describe important information to contextualize the data source, including:		
	1.1. type of care setting		
	1.2. geographical location.		
	2. Describe all relevant study period dates, including periods of recruitment, exposure, follow-up, and data collection.		
	3. Clearly identify missing data components in the data collection.		
	4. For studies that propose the use of a data source from a country other than Canada, provide:		
	4.1. a rationale for selecting the data source		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	4.2. an explanation of how these factors might affect the generalizability of the study results to the population in Canada		
	4.3. background information about the health care system		
	4.4. a description of prescribing and utilization practices		
	4.5. information on the use and market availability of the intervention and comparators of interest throughout the study period.		
Section 3: Data specifications — access, cleaning methods, and linkage	1. Describe the extent to which the investigators had access to the database population used to create the study population and major aspects of data provenance.		
	2. Provide information on the data-cleaning methods used in the study. Share any data-cleaning code leveraged. If not provided, justify.		
	3. Report whether data were organized by a Common Data Model structure.		
	4. Describe the usage of data and consent for data sharing. Provide consent documents, if relevant.		
	5. Describe data collection methods.		
	6. Quality of the data and relevant metrics to assess the data quality should be reported.		
	7. Describe any variability between data sources and the impact of changes over time in the data.		
	8. Describe if any data linkage was conducted and the methods used for the linkage.		
	9. Report who (e.g., which organization) performed the data linkage, if applicable.		
	10. Describe the performance characteristics of the data linkage and the number of individuals linked at each stage of linkage.		
Section 4: Data sources, data dictionary, and variables	1. Provide and describe all data sources, including the specific version and date of the last update of the database.		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	2. Describe the characteristics of the health setting and context of data collection.		
	3. Describe details of data continuity and completeness.		
	4. Include the names, dates, and/or version numbers of when data were extracted for research use by the data vendor or organization.		
	5. Include the search and/or extraction criteria applied if the source data are a subset of the data from the vendor or organization, and provide calendar date ranges.		
	6. Provide source(s) of data for each variable of interest.		
	7. Describe how variables of interest were measured and if they have been adjudicated or validated in the population of interest.		
	8. Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable.		
	9. Specify definitions and lookback windows for all variables.		
	10. Report whether any variables could be time-varying (e.g., how the variable could change over time and when it was redefined in relation to time-varying exposures).		
	11. Report important variables that could not be captured and their anticipated impact on study results.		
	12. Provide information on deviations from the a priori protocol in variable measurement.		
Section 5: Participants	1. Provide inclusion criteria used to identify the study population.		
	2. Justify exclusion criteria and how they may affect the overall interpretation of the research.		
	3. Describe study population characteristics relative to the target population in Canada.		
	4. Provide all codes or algorithms used to define the inclusion and exclusion criteria, where possible.		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	5. Specify the time period (e.g., lookback window) over which inclusion and exclusion criteria were assessed.		
	6. Recommendations for specific study designs:		
	6.1. For cohort studies, provide details leading to the analyzed cohort, including definitions for exposure groups, cohort entry and end dates, matching criteria, and censoring/ follow-up.		
	6.2. For prospective cohort studies, describe recruitment processes.		
	6.3. For case-control and case-crossover studies, provide details of case and control ascertainment, the source population for nested studies, sampling methods, and matching criteria.		
	7. Report the numbers of participants at each stage of the study and reasons for nonparticipation. Consider illustrating this information using a flow diagram.		
	8. Provide characteristics of study participants. If not available or feasible, explain why.		
	9. Indicate missing data for each variable of interest.		
	10. Compare treatment or exposure groups.		
	11. Specify the number of participants included in each analysis and the analysis strategy (e.g., per-protocol, ITT) and provide details on the number or proportion of subjects excluded from each analysis, and the reasons for exclusion.		
Section 6: Exposure definitions and comparators	1. Define the requirements for the exposure definition (e.g., single, multiple, or continuous exposure) and relevant start and stop windows for assessing exposures.		
	2. Specify the data source(s) from which exposure information was obtained, including validity and any limitations in exposure measurement.		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	3. Specify the exposure-outcome risk window and discuss how it aligns with the known or anticipated relationship between the exposure and outcome timing.		
	4. If no comparator was used, justify why not.		
	5. Define the comparator group(s) (e.g., active comparator, historical comparator).		
	6. Provide justification for the comparator used, including potential implications on the study findings.		
	7. Discuss any changes in patterns of use of the exposure and comparator(s) over time and how they may affect the results. Report any methods used to adjust for these changes.		
	8. Specify how adaptations to the intervention and/or comparator were permitted and recorded.		
Section 7: Outcomes	1. Report definitions for all study outcomes (primary, secondary, and exploratory), where possible.		
	2. Provide a rationale for the outcomes studied and discuss relevant outcomes not included in the study. Consider the use of a core outcome set if one is available for the condition of interest under study.		
	3. Provide information about the validity of all outcome definitions.		
	4. Describe whether the timing of the outcome can be accurately measured.		
	5. Specify whether the outcome studied is a surrogate measure of a clinical (patient-centred) outcome and, if so, the strength of the relationship between the surrogate outcome and major clinical outcome(s) of interest.		
	6. Discuss whether outcome misclassification could occur between treatment groups.		
	7. Report whether a control outcome was used and justify the control outcome(s) selected.		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
Section 8: Bias, confounding, and effect modifiers or subgroups effects	1. Report all procedures used to address potential sources of bias.		
	2. Specify how potential sources of bias could influence the outcomes of the analyses.		
	3. Specify variables that were considered known or potential confounders in the analysis.		
	4. Describe how confounder variables were selected and if they were informed by a causal diagram.		
	5. Describe and compare the distribution of measured baseline confounding variables between treatment groups.		
	6. Report whether any potential confounders could not be measured and specify the anticipated impact of these confounders on study results.		
	7. Report whether time-varying confounding was considered and, if not, justify why not.		
	8. Specify the methods used to conduct sensitivity analyses that test key assumptions and limitations of the data and, if no sensitivity analyses were conducted, explain why not.		
	9. Specify known or potential effect modifiers.		
	10. Describe any effect modification or subgroup analyses that were conducted and if they were specified a priori. Include if they were identified and conducted based on prespecified rationale such as previous studies or biological rationale. If no effect modification or subgroup analyses were used, justify why they were not needed.		
	11. If effect modification or subgroup analyses were used, describe the methods and present separate results for each subgroup.		
Section 9: Statistical methods	1. Indicate the software used for the statistical analysis, including software package, version, and analytic tools employed (e.g., macros).		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	2. Provide access to the statistical code used or, if the code cannot be shared, explain why not.		
	3. Report all statistical methods used and justify their selection, including, as applicable:		
	3.1. all variables included in regression models		
	3.2. the method of variable selection for regression models		
	3.3. methods used to control for confounding		
	3.4. methods used to account for missing data		
	3.5. how follow-up time and changes in exposures were handled		
	3.6. subgroup analyses and effect modification		
	3.7. as applicable: stratification, propensity score estimation and assumptions, meta-analysis methods, validity of instrumental variables.		
	4. Quantify the precision of all estimates using confidence intervals.		
	5. Report the threshold of statistical significance used.		
Section 10: Study findings	1. Summarize key results (estimated effect measures, measures of precision) with reference to each study objective and/or hypothesis for primary and secondary outcomes, and delineate these results by each treatment or exposure group.		
	2. Provide numbers of outcome events or summary measures of outcomes (or exposures in case-control studies).		
	3. Report both absolute and relative effect measures for binary outcomes, including their measure of precision.		
	4. Report category boundaries when continuous variables are categorized, and consider translating estimates of relative risk into absolute risk.		

Section	Checklist item	Reported on page number(s)	If not reported or not applicable, justify why
	5. Report unadjusted and adjusted estimates, including their measure of precision and confounders used for adjustment.		
	6. Report other prespecified analyses conducted (e.g., subgroup analyses, interactions, sensitivity analyses).		
	7. Describe any unplanned analyses performed secondarily (not defined a priori) and indicate these as exploratory.		
	8. Avoid selective reporting of results.		
Section 11: Interpretation and generalizability	1. Provide an interpretation of the primary and secondary study results, as applicable.		
	2. Interpret the findings from adjusted and unadjusted results, as applicable.		
	3. Discuss the precision of the effect measure(s).		
	4. Discuss how potential biases and sensitivity of study assumptions may impact the results and subsequent interpretation.		
	5. Discuss the implication of findings for clinical practice, including the risk-benefit profile of the treatment, if applicable.		
	6. Interpret study findings in relation to current literature.		
	7. Discuss the generalizability (external validity) of study results to the population in Canada.		
Section 12: Limitations	1. Provide a consideration of limitations of the study, including the data source, missing data, bias and confounding, imprecision or sample size limitations, and whether results are clinically meaningful.		
	2. Discuss the plausibility of results and whether results could be due solely to bias, chance, or confounding.		