RWE Guidance
CADTH Boilerplate Title page and acknowledgements

Authorship list

Authorship Team

Mina Tadrous, Lead- University of Toronto
Kaley Hayes- Brown University
Theresa Aves- University of Toronto
Christine Fahim- Unity Health Toronto

Expert Panel

Canadian members

Robert W. Platt -McGill
Erin Strumpf -McGill
Lisa M. Lix- University of Manitoba
Ted McDonald - UNB
Kelvin Chan - CanREValue
Scott Klarenbach – University of Alberta
Jason Guertin - Université Laval
Jeff Round - IHE, University of Alberta
Sanja Stanojevic – Dalhousie University
Mary De Vera - UBC

International members

Shirley Wang - Harvard
Daniel Prieto-Alhambra - Oxford (NDORMS)
Seamus Kent - NICE
Dalia Dawoud - NICE
Donna Riveria - FDA

Stakeholder Observers and Experts

Erik Dorff- Statistics Canada
Andrew Raven- Health Canada
Catherine Njue- Health Canada
Geneviève Plamondon- INESSS
Sara Beha- INESSS
Tom Samaha- INESSS
Deborah Cohen- CIHI
Nicole Mittmann- CADTH
Sheri Pohar- CADTH
Hongbo Yuan- CADTH
Matthew Bryan- CADTH
Amanda Allard- CADTH
Karen Lee- CADTH
Cody Black- CADTH
Mike Innes- CADTH

Abbreviations

HTA- Health Technology Assessment

RWD- Real World Data

RWE- Real World Evidence

Summary

Real-world evidence (RWE) is clinical evidence on the use, safety, effectiveness and cost of medical technologies that is derived from real-world data (RWD).[1] 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.[1-4] However, as the volume and types of RWE have rapidly expanded, there is a need for standardization in the quality of reporting and methodology for HTA and regulatory submissions involving RWE.[1]

The variety and complexity of RWD sources, study designs, and analytical methods make the evaluation of submissions that include RWE challenging. Therefore, in an effort to optimize the utility and transparency of RWE submissions to HTA and regulatory bodies, it is important to develop a set of methodological considerations and core standards for reporting.. Recent global initiatives have focused on developing tools to improve reporting, transparency and reproducibility of RWE studies. However, further guidance is needed to ensure adequate consideration of methodological issues relevance and standards for reporting within the Canadian context in order to facilitate the use of RWE submissions for regulatory and reimbursement decisions. Additionally, other reporting guidance documents have largely aimed to establish standards for reporting for either HTA or regulatory submissions, not both types of submissions. To address these needs, this document serves as a comprehensive, fit-for-purpose, and credible reporting guidance that aims to harmonize current RWE submission principles for Canadian HTA agencies and regulators while maintaining alignment with international standards.

What is RWE?

Randomized controlled trials (RCTs) are the gold standard for establishing the efficacy and safety of health technologies. However, trials often produce results for specific populations and settings within controlled environments, limiting the generalizability of results to patients in a real-world setting. Additionally, in some circumstances, such as the evaluation of treatments for rare diseases, large trials that are sufficiently powered to detect the effects of interventions on important clinical outcomes are not always feasible. RWE can provide direct and generalizable evidence that fills knowledge gaps left by RCTs 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 information found in a variety of sources such as electronic medical records, clinical and disease registries, and administrative databases.[5,6] RWD can also be drawn from other prospective study designs including cohorts, registries and clinical and pragmatic trials. RWD can provide information about ’ medical history, demographics, socioeconomic factors, health behaviors, experiences, clinical and functional outcomes, resource use and costs.

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, underserved and understudied populations, or patients with a high burden of multimorbidity. RWE about these populations can inform decision-making about real-world effectiveness, safety and patient experience on a population level through the use of expanded sample sizes and longer follow-up periods that may not be feasible in clinical trials. RWE can also offer insights into healthcare

providers' and patients' perspectives on issues related to accessibility, acceptability, preferences, , and ease of use of health technologies. Furthermore, depending on the design, RWE studies may be able to provide evidence in a more cost- and time-efficient manner than clinical trials. There is potential to leverage RWE across phases of the health technology development lifecycle; for example, RWE 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 generate evidence in all scenarios. RWE may be subject to bias and confounding.. Issues with RWE can include non-random treatment assignment and unblinded ascertainment of outcomes that may not be adjudicated and verified to the same degree as in clinical trials). Further, RWE can be complex and vary largely in quality[1]; thus, reaching appropriate conclusions requires transparent reporting and careful interpretation. Clear standards are needed to guide the reporting of RWE for decision-making.[1]

Overall Purpose:

The purpose of this guidance is to standardize the reporting of RWE submissions by providing best practices for reporting and highlighting important methodological considerations for those undertaking and submitting RWE studies of healthcare technologies in Canada.

Specific Objectives:

1. Identify existing global guidance on RWE reporting and standards through an environmental scan to create an initial draft of Canadian RWE submission standards.
2. Establish consensus on items to be included in the core reporting standards for Canadian RWE submissions through engagement with national and international experts in RWD and RWE.

About this Guidance

Since standards for generating RWE have already been developed on a global level (see Appendix), this document was created to best align with those of international initiatives in this space, with strong consideration of the Canadian context. The aim is for all RWE submissions 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 submissions.

Of note, this document does not offer guidance as to *when* or *why* RWE should be used (e.g., whether an RWE study is appropriate for a particular research question). This is a critical step for future initiatives but first requires a strong foundational guidance on the overall reporting of RWE. Moreover, RWE does not and should not serve as a replacement for clinical trials, but must supplement existing trial evidence and should be considered as a part of the broader body of evidence for decision-making. The present guidance aims to ensure that each submission will transparently provide Canadian regulators and HTA agencies with the information they require to interpret the submission and 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.

Lastly, this guidance will be a living document, such that it will require updates, revisions, and extensions over time. This guidance document is meant to allow sufficient flexibility for submissions to accommodate the heterogeneous nature of RWE and its rapid evolvement while ensuring that submissions are sufficiently detailed and transparent to facilitate regulator and HTA decision-making.

Background and Methods

This document was developed through an iterative process with the support of Canadian and international RWE experts and stakeholders (see authorship list). The Appendix of this document contains full methodological details of how this guidance was developed. In brief, an environmental scan on the use of RWE and RWD to Support Decision-Making in Drug Assessments[7] was leveraged and expanded to identify preliminary relevant documents for review. Additional documents were identified using a citation-review method and expert consultation. The resulting set of documents including international RWE guidance, systematic reviews, reporting guidelines, and policy statements were reviewed to develop candidate reporting recommendations for expert review. In total, 37 documents were reviewed (See Appendix for the full list). Data on 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 authorship team that was shared among the group of 15 national and international experts. In this survey, experts were asked if each candidate recommendation should be included in a guidance for standards on the reporting of all RWE submissions intended for any health-technology or regulatory use in Canada. A recommendation was included or excluded based on whether there was ≥70% consensus on its importance by the experts. Recommendations not achieving this level of consensus were discussed in a large group meeting and revised as needed. A draft report was shared with the experts for review and feedback which was subsequently collated and incorporated by the authorship team. Outstanding points of disagreement were discussed at a second in-person meeting and consensus was reached at the same level as previously defined.

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 engagement process and public comment 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 review and feedback for 4 weeks. The authorship team also held multiple in-person and virtual events during the stakeholder feedback period (See website for a full list of events) and leveraged established networks to provide opportunities for additional feedback to be submitted. The authorship team reviewed the feedback and incorporated it as required; all modifications were reviewed by the expert panel for final approval.

Overview and Structure

The guidance is reported in 14 sections. In each section, we present an overview, a narrative of the recommendations in detail, followed by a summary list of recommendations at the end of each section. These sections may have overlapping concepts, but each has a specific goal and purpose. **The summary lists of recommendations and the checklist for submissions are not intended to replace a careful review of the text that contains critical information needed to develop an adequate submission.**

Recommended reporting guidance is provided for the following study components:

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

Section 1: Research Questions and Study Design

Overview:

Study reporting should aim for full transparency to allow for straightforward interpretation of study design choices and to facilitate potential reproducibility. 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. Thus, central and foremost to any submission is full reporting on the study question and study design. Reviewers of the submission must be able to understand the main purpose and aims of the study to interpret and appraise the reporting of the study and its design throughout the submission.

Specific Considerations and Recommendations

Study Aim and Question:

The aim and research 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 research study. The research question should be phrased by using the PICOTS template: Population, Intervention, Comparator, Outcome, Timing and Setting.[8] This template should be adapted for varying study designs depending on the research question being asked. Each of these components should be reported in a precise 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, initial primary and secondary outcomes and the main measure(s) of effect (e.g., hazard ratios) should also be reported, as they likely influence 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 use of matching and the overall allocation and matching criteria (e.g., 1 to 1 hard-matching based on age and sex). It is suggested that reporting be aligned with the standards of established reporting for the type of study design employed.[9-12]

To improve transparency and ensure the interpretability of any submission, the use of study design diagrams is suggested.[13] These diagrams illustrate important components of the study design, including observation windows, exposure windows, covariate measurement periods, washout periods, and lag periods. 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.[13]

Other Recommendations and Additional Transparency:

Development and registration of an a priori protocol prior to conducting the study is recommended. The protocol should be registered to a permanent platform, that assigns a unique study identifier, and is maintained by a third party. If a study protocol was developed, it should be referenced in the initial reporting of the study design, including a reference to the

protocol and the registration number.[13] In the submission, 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 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.

It is suggested to include a description of each team member involved in the study, specifying their role, organizational affiliation, education, title, and experience. Inclusion of patient partners is encouraged, and their involvement should be clearly described. 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. Lastly, reporting of study governance, especially if multiple partners are involved, is recommended to allow for full transparency in study structure and execution. Study governance reporting must include all sources of funding and potential conflicts of interest for external groups involved in the submission, 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

1. Report a clearly stated aim and research 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. Encourage the development of and reference to a priori protocols
8. Describe all study team members including the role of patient partners
9. Describe the study governance structure, especially who was responsible for final decision-making
10. Report any research ethics approvals (or equivalent)
11. Disclose sources of funding

Section 2: Setting and Context

Overview:

Given the global nature of RWE and the potential to leverage data from multiple jurisdictions, it is essential that all reporting include detailed information on the study setting and context (i.e., health system factors like universal healthcare), even if the data come from a Canadian setting. This reporting is especially critical for data sets that are not regularly used by various research groups or are novel (e.g., single system electronic health record data or novel registries). Even when well-known or common datasets (e.g., the Clinical Practice Research Datalink or ICES databases) are used, detailed reporting should be provided, as data sources change over time. Non-Canadian data can be an acceptable source of RWE but must have important components reported in order for reviewers to understand the RWD/RWE's strengths, limitations, and generalizability to the Canadian context. The reporting components in this section ensure that reviewers of the submission can understand the implications of the data source, 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); thus, 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), healthcare system context, and time periods in which data were available. Reporting on the healthcare system context should include: the overall healthcare 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 research 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 for submissions, but given the importance of generalizability, studies leveraging non-Canadian data must have important components reported. Foremost, the rationale for why Canadian data was not included should be described; Justification for use of this data source and its alignment with the study aims and research questions must be articulated. In addition, an explanation of how all the study setting factors might affect the generalizability of the results to the Canadian population must be provided. Background information about the healthcare 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.

Finally, information on the market availability of the intervention and comparators of interest throughout the study period (e.g., regulatory approval dates, formulary restrictions) should be reported, highlighting if and how they differ from the Canadian market.

Section 2: Summary of Recommendations

12. Describe important information to contextualize the data source, including:
    a. Type of care setting
    b. Geographical location
13. Describe all relevant study period dates, including periods of recruitment, exposure, follow-up, and data collection
14. Report the restrictions and limitations on codes to those identified from certain settings
15. Clearly identify missing areas of insight in your data collection
16. For studies that propose the use of a data source from a country other than Canada, provide:
    a. A rationale for selecting the data source
    b. An explanation of how all these factors might affect the generalizability of the study results to the Canadian population
    c. Background information about the healthcare system
    d. A description of prescribing and utilization practices
    e. Information on the market availability

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.[14-15] Reporting on data provenance is important to ensure credibility of the data leveraged and full transparency of data specifications. It is recognizable that the current landscape of data access is complex, with various nuances related to data ownership, privacy regulations, and intellectual priority. Given the importance of transparency for submissions 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 submission and should be highlighted as a limitation(s).

Specific Considerations and Recommendations

Data Access:

Describe the extent to which the investigators had access to the data. Describe the usage of data and consent for data sharing (and with whom). Data ownership and processes for access must be described including whether a data vendor or organization was used, and any potential costs associated with data access. These statements cover registry data access with interested parties from other countries and/or international organizations. 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, out-of-range values, 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 unable to report these components for the entire data source, it is recommended to perform these steps for the analytical study data at a minimum.

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.[15] 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.[15] Any variability between data sources and the impact of changes over time in the data should be reported (e.g., pre- vs. post-COVID-19 period).

Data Linkage:

State whether the study included person-level, institutional-level, or other levels of linkage across databases. Report if consent was required for linkage and how it was attained. Describe the methods of linkage, including whether the linkage was deterministic or probabilistic, 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 unlinked, changes in linkage performance over time). If available, include the number of individuals with linked data at each stage if a multi-stage approach was used to better understand how representative the final study data is of the population of interest.

Other Recommendations:

Any methods 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 (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

17. Describe the extent to which the investigators had access to the database population used to create the study population
18. Provide information on the data-cleaning methods used in the study. Share any data cleaning code leveraged. If not provided, justify
19. Describe the usage of data and consent for data sharing. Provide consent documents if relevant
20. Describe data collection methods
21. Quality of the data and relevant metrics to assess the data quality should be reported
22. Describe any variability between data sources and the impact of changes over time in the data
23. Describe if any data linkage was conducted and the methods used for the linkage
24. Report who (e.g., organization) performed the linkage
25. Describe the performance characteristics of data linkage and the number of individuals linked

Section 4: Data Sources

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., Canadian Institute for Health Information, ICES). 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 validation status, as available, are needed to provide insight to applicability and validity.

Specific Considerations and Recommendations

Data Sources and Context:

Indicate all sources of data being used in the research, including how they were obtained. 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 using administrative claims data or registries, reporting the data completeness (i.e., continuity of coverage) is important, as individuals often enroll and disenroll in different health plans in relation to changes in employment or other life circumstances. Any discontinuities or variations among source data (e.g., intra-patient or inter-provincial differences in data availability) should be documented. Specify the source(s) of data for each major variable of interest in the study (detailed further in Section 8).

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, it is necessary to request it from the data vendor for inclusion in the submission.

Section 4: Summary of Recommendations

26. Provide and describe all data sources
27. Describe the characteristics of the health setting and context of data collection
28. Describe details of data continuity and completeness
29. Include the names, dates and/or version numbers of when data were extracted for research use by the data vendor or organization
30. 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
31. Provide source(s) of data for each variable of interest
32. Describe how variables of interest were measured and if they have been adjudicated or validated in the population of interest

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. All methods and decisions that led to the final, analyzed study population (e.g., random sampling from the source population, exclusion criteria) should be described in a stepwise manner, with definitions provided (e.g., exposure groups, cases, controls). Ideally, these steps are 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). Investigators must clearly describe all inclusion and exclusion criteria used to identify the study population along with detailed justification for each exclusion criterion.

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 enrollment 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 enrollment gaps were allowed prior to 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 (e.g., 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 of 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/gender, race/ethnicity, and other characteristics important to consider for diversity according to up-to-date guidance[16] and should emphasize how the other historically underrepresented groups in research are included to the extent that is possible.[16] All codes or algorithms (e.g. drug, diagnosis, procedure, lab codes, etc.) that were used to define the inclusion and exclusion criteria should be specified in the submission. 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 were excluded from the study, investigators should justify this approach, providing a detailed explanation of the exclusions, 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, 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[17], prevalent-user, or other type of cohort study design was used. Where possible, a new-user design is preferred because follow-up for all patients begins at the same time point in the treatment course. In contrast, the prevalent user design often begins follow-up for patients at different time points in their treatment course, potentially resulting in issues such as depletion of susceptibles.[13,14] If a new-user design was chosen, specify the lookback period 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.

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 prior to 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(s) 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.[18]

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 implicate 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. If the case-control study is nested, describe the cohort or source population from which it is derived as discussed above. 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 callipers, and hard-matched covariates between cases and controls (e.g., sex).

Section 5: Summary of Recommendations

38. Provide inclusion criteria used to identify the study population
39. Justify exclusion criteria and how they may affect the overall interpretation of the research
40. Describe study population characteristics relative to the target Canadian population
41. Provide all codes or algorithms used to define the inclusion and exclusion criteria
42. Specify the time period (e.g., lookback window) over which inclusion and exclusion criteria were assessed
43. Recommendations for specific study designs:
    a. 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
    b. For prospective cohort studies, describe recruitment processes
    c. For case-control and case-crossover studies, provide details of case and control ascertainment, the source population for nested studies, sampling methods, an matching criteria

552
553

Section 6: Participant Characteristics

Overview:

A detailed description of study participant characteristics is critical for assessing potential confounders/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(s) 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 non-participation. It is recommended to use a visual aid or figure to represent the reported number of participants at each stage of the study. Statistical comparisons of participant characteristics between treatment or exposure groups must be fully described along with a description of the extent and handling of missing data as well as the treatment estimate.

Specific Considerations and Recommendations

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).[11,19] Consider illustrating this information using a diagram to report the flow of participants throughout the study. At each stage, investigators should provide reasons for non-participation and exclusion. Investigators should provide a breakdown (e.g. 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).

Disposition of Participants:

Provide comparisons of participant characteristics by treatment or exposure groups. Indicate the number of participants with missing data for each variable (characteristic) of interest. Include the number of participants in each analysis conducted and the analysis strategy that was used. For example, report whether the analysis was conducted according to the original exposure groups (e.g., ITT) or whether a person’s exposure group could change over follow-up. 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.[20-22]

Section 6: Summary of Recommendations
39. Report the numbers of participants at each stage of the study and reasons for non participation. Consider illustrating this information using a flow diagram 40. Provide characteristics of study participants 41. Indicate missing data for each variable of interest 42. Compare treatment or exposure groups 43. Specify the number of participants included in each analysis and the analysis strategy (e.g. per-protocol, ITT)

604
605

Section 7: Exposure Definitions and Comparators

Overview:

A critical element of RWE submissions is clear and justified definitions of the exposures and comparators used in the study.[23] The term "exposure" can refer to a host of treatments and factors, including drugs, devices, or clinical conditions. In this section, we focus 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 comparator group (or control period for self-controlled studies where 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, investigators must specify how adaptations to the intervention and/or comparator were permitted and recorded. 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). Specify the data source(s) from which exposure information was obtained including 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. If no comparator was used in the study, explain why. Justify the potential implications of the comparator group if it does not specifically include:

1) an active comparator,
2) a drug used to treat the same disease,
3) patients reasonably expected to have the same level of disease severity, AND
4) patients from the same time period as the exposed cohort.

If an external (e.g., historical) 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 (e.g., similar clinical guidelines for the disease of interest). Similarly, if the comparator group is from the same individual but a time period prior to the exposure (e.g., in a self-controlled case series), 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 of the comparator group compared to the treatment group.

Other Exposure and Comparator Considerations:

Discuss any changes in patterns of use of the exposure and comparator over time and how they may impact the study findings. Describe how exposure switching 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).

Section 7: Summary of Recommendations
44. Define the requirements for the exposure definition (e.g., single, multiple, or continuous exposure) 45. Specify the data source(s) from which exposure information was obtained, including any limitations in exposure measurement 46. Specify the exposure-outcome risk window and discuss how it aligns with the known or anticipated relationship between the exposure and outcome timing 47. If no comparator was used, justify why not 48. Define the comparator group(s) (e.g., active comparator, historical comparator) 49. Provide justification for the comparator used including potential implications on the study findings 50. Discuss any changes in patterns of use of the exposure and comparator(s) over time and how it may affect the results 51. Specify how adaptations to the intervention and/or comparator were permitted and recorded

Section 8: Outcomes

Overview:

The utility of RWE for decision-making relies heavily on whether the outcomes studied are relevant to the research question being asked. 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. This section must contain detailed information on:

1. Study outcomes and their definitions,
2. References on the validity of these outcome definitions (including the strength of association between any surrogate outcomes and clinical outcomes, if applicable and known),
3. A discussion of the relevance of study outcomes to real-world practice,
4. 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 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 expected to occur too rarely or have too unclear of a relationship to the exposure to show effects or make conclusions, but are included to explore future hypotheses.[24] Report which outcomes were specified *a priori* versus which were *post hoc*. Report all changes to the planned protocol and reasoning. Explain why the study outcomes were selected (e.g., relevance to clinical practice, safety concern, patient/caregiver consultation) and discuss any relevant outcomes that were not studied with a justification as to why they were not included.

Specify the definitions used for all study outcomes.[25] 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. 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, secondary, any position), as applicable. If validation studies of the codes or algorithms were conducted (i.e., studies that estimate the sensitivity, specificity, etc. of the code or algorithm), reference these studies and include 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/discontinuation of the exposure(s). For example, outcome ascertainment of a myocardial infarction is likely relatively precise versus 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., internal study), provide detailed methods and results from this validation study - ideally with sensitivity, specificity, and positive predictive values of the outcome definition. 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 9th 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.[26] If a validated instrument was not used to capture the outcome, justify why not. Discuss whether the outcome or its measurement may be subject to clinical judgment (e.g., the outcome is an opinion on whether the patient's condition has improved). If applicable, 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 outcomes were assessed or validated on the individual case level (e.g., through record review by a specialist blinded to the exposure(s) under study, in order 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 (final status is "possible," "uncertain," etc.).

Outcome Selection and Surrogate Outcomes:

Drawing causal conclusions from RWE is discouraged; however, discussing selected outcomes relative to their location on the causal pathway is recommended. When available, clinical outcomes (e.g., major cardiovascular events) are preferred to surrogate outcomes (e.g., changes in laboratory values) 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). 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 endpoints if a surrogate outcome was used as the primary endpoint.

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 control outcome(s) was used. A control outcome in this context is an outcome that is not expected to be related to the intervention/exposure.[27] Therefore, the control outcome

should not be associated with the intervention/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(s) of interest.

Section 8: Summary of Recommendations

52. Report definitions for all study outcomes (primary, secondary, and exploratory)
53. Provide a rationale for the outcomes studied and discuss relevant outcomes not included in the study
54. Provide information about the validity of all outcome definitions
55. Describe whether the timing of the outcome can be accurately measured
56. Specify whether the outcome studied is a surrogate measure of a clinical (patient centered) outcome and, if so, the strength of the relationship between the surrogate outcome and major clinical outcome(s) of interest
57. Discuss whether outcome misclassification could occur between treatment groups
58. Report whether a control outcome(s) was used and justify the control outcome(s) selected

Section 9: Variables and Data Dictionary

Overview:

Just as detailed exposure and outcome definitions are critical for interpretation of the study, so are details on the methods used to clean data and define all other variables. Variables other than the exposure and outcome include 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

Types of Variables and Measurement:

The primary variables of interest for RWE studies outside of the exposure and outcome are confounders and effect modifiers. The definitions and types of confounders and effect modifiers are discussed in more detail in Section 10. Provide a detailed definition of each variable, following similar principles outlined in Sections 6 (Exposures) and 7 (Outcomes); further details are provided in the next section (Data Dictionary). These definitions must specify the lookback window used to ascertain variables (e.g., in the 365 days prior to the date of first exposure). 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 re-defined in relation to time-varying exposures. Reporting requirements for predictive modeling and mediation analyses are outside of the scope of this document, but predictor variables and mediators (intermediaries) should have similar reporting provided as for potential confounding variables.

Data Dictionary:

Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable. Include information on how types of data for variables were collected; timing of capture, including the lookback window; the source (e.g., clinical diagnoses, tests, procedures, prescriptions). It is important that naming and variable definitions remain consistent throughout the submission. 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 were used, report any differences in how data were coded, recorded, or collected between sources. 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.

Section 9: Summary of Recommendations
59. Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable 60. Specify definitions and lookback windows for all variables 61. Report whether any variables could be time-varying (e.g., how the variable could change over time and when it was re-defined in relation to time-varying exposures) 62. Report important variables that could not be captured and their anticipated impact on study results 63. Provide information on deviations from the *a priori* protocol

832

833

834

Section 10: Bias, Confounding, and Effect Modifiers/Subgroup Effects

Overview:

Bias is systematic error that results in an incorrect estimate of the association between the exposure and the outcome. Unlike randomized trials, in RWE studies, the treatment assignment is not controlled. RWE is therefore often subject to confounding and other biases related to non-random 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 submissions 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.[28,29] In brief, the main types of bias are information bias, selection bias, and bias from confounding. Information bias arises when key study variables (e.g., exposure or outcome) are measured differentially between treatment groups. For example, immortal time bias is common in RWE and occurs when follow-up time is included during which the study outcome cannot occur.[30] 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 study outcomes. Confounding by indication is a particularly important concern when comparing medication and device exposures.[31]

This section requires a critical review and reporting of assumptions. 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 or analysis) helped to mitigate or account for this bias. Explicitly report if there is a risk of a bias but it was not able to be addressed. Regardless of whether methods were employed to mitigate or account for a bias, investigators must explicitly discuss how results would be impacted by each identified bias (magnitude and direction of effect towards or away from the null value).

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 richness of RWE and often large study populations, effect modification may be explored in most studies 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, race/ethnicity) and any established effect modifiers from the literature should be explored.

Specific Considerations and Recommendations

Bias:

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, informative censoring, and non-response rates between treatment groups and potential implications. Sensitivity analyses used to test *specific* assumptions and potential biases are described in more detail in Section 10.

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 and the anticipated impact of these confounders on results. 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. 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).[33] 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 compare their baseline equivalence. Discuss whether time-varying confounding was considered, especially if participants could switch between the treatment and comparator groups.

Sensitivity Analyses:

Sensitivity analyses test key assumptions on which primary results are based (e.g., whether some analyses were reported to evaluate the potential for a biased assessment of exposure or outcome, such as analyses where the impact of varying exposure and/or outcome definitions was tested to examine the impact on results). 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. Report all methods used for sensitivity analyses and specify whether each sensitivity analysis was 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. If effect modification/subgroup analyses by main demographic variables (e.g., age, sex, and race) or pertinent co-morbidities were not examined, explain why not. If subgroup analyses are employed, describe the methods used to examine the subgroups and interactions. Present effect measures for separate subgroups defined by the effect modifiers.

Section 10: Summary of Recommendations

64. Report all procedures used to address potential sources of bias
65. Specify how potential sources of bias could influence the outcomes of the analyses
66. Specify variables that were considered known or potential confounders in the analysis
67. Describe how confounder variables were selected and if they were informed by a causal diagram
68. Describe and compare the distribution of measured baseline confounding variables between treatment groups
69. Report whether any potential confounders could not be measured and specify the anticipated impact of these confounders on study results
70. Report whether time-varying confounding was considered and if not, justify why not
71. 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
72. Specify known or potential effect modifiers
73. Describe any effect modification or subgroup analyses that were conducted and if they were specified *a priori*. If effect modification/subgroup analyses by main demographic variables (age, sex, race/ethnicity) or by pertinent comorbidities were not employed, justify why these analyses were not needed
74. If effect modification/subgroup analyses by main demographic variables (age, sex, race/ethnicity) or by pertinent comorbidities were employed, describe the methods used and present separate results for each subgroup

Section 11: Statistical Methods

Overview:

The choice of statistical methods and inherent modeling assumptions can greatly impact study findings; however, statistical methods are often underreported in RWE. For statistical reporting, transparency in the methods used to generate results is critical. Provision of all or at least part of the code used is one straightforward method to providing transparency. If code 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 endpoints, the estimand principle should be applied as much as possible.[34] The statistical methods used should ideally be decided a priori in order to ensure that they are not overly data-driven. Any changes to the statistical methods should be documented prior to the analyses being conducted, and reasons for the change should be documented. Approaches for handling clinically relevant intercurrent events 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 critical than estimates of statistical significance alone for interpreting study results.

Specific Considerations and Recommendations

Essential Statistical Reporting:

Indicate the software(s) used for the statistical analyses including software package, version, settings, packages or analytic procedures. 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 studies utilizing regression analysis, it is essential to identify if the study identified which variables were included and how they relate to the outcome. Describe methods used for variable selection (e.g., e.g., criteria for stepwise selection of variables in multivariable models). Detail any statistical methods used to control for confounding and to account for missing data, if applicable. Report how multiplicity (i.e., multiple testing) was handled, particularly for outcomes that were not specified *a priori*.

Method-Specific Statistical Reporting, as applicable:

Describe methods used to account for differential follow-up time between exposure groups. Describe 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), as applicable. Instrumental variables should be used with caution given the strong assumptions required for this method. [35,36] If an instrumental variable (is used, report methods used to assess the validity of the instrument.[35-36] 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 study findings.[37] Instead, estimates of precision should be quantified (e.g., via confidence intervals). Specifically, authors should not describe results as "statistically significant" or "non-significant" or rely on thresholds for p values, but rather report the exact p value together with an estimate of precision like a confidence interval.

Section 11: Summary of Recommendations

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

Section 12: Study Findings

Overview:

Central to any RWE submission 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 each exposure group for all primary and secondary outcomes. 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. Additionally, these reported values should be accompanied by results from subgroup and sensitivity analyses, and interactions. Unplanned analyses, if conducted, must be clearly presented as post-hoc.

Specific Considerations and Recommendations

Reporting Main Analyses:

Outcomes should be reported in the manner and order by which they were presented in the methods. It is imperative that the outcomes of a study are presented in an objective manner, 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. All primary and secondary outcomes 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 needed. Methods that account for differential follow-up (e.g., time-to-event regression models) between subjects should be used when applicable. When time-to-event analyses are used, consider translating estimates or relative risk to absolute risk for a meaningful time period (e.g., 1-year mortality risk). Report both unadjusted and adjusted estimates, including their measure of precision and the confounders used for adjustment.

Reporting Other Analyses:

All other pre-specified 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. Selective reporting of results is not best practice.[26]

Section 12: Summary of Recommendations

80. 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
81. Provide numbers of outcome events or summary measures of outcomes (or exposures in case-control studies)
82. Report both absolute and relative effect measures for binary outcomes including their measure of precision
83. Report category boundaries when continuous variables are categorized and consider translating estimates of relative risk into absolute risk
84. Report unadjusted and adjusted estimates including their measure of precision and confounders used for adjustment
85. Report other pre-specified analyses conducted (e.g., subgroup analyses, interactions, sensitivity analyses)
86. Describe any unplanned analyses performed secondarily (not defined *a priori*) and indicate these as exploratory
87. Avoid selective reporting of results

Section 13: Interpretation and Generalizability

Overview:

A thoughtful and balanced interpretation of study results is critical to any RWE submission. Primary and secondary 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. Though 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, patient and/or caregiver involvement should be considered to support the interpretation and generalizability of study findings to a Canadian context.[38]

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). Re-iterate potential biases that could affect study results to allow for appropriate clinical interpretation of findings. 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 Canadian population, considerations of equity and diversity of participants[16] as well as inclusion and exclusion criteria. Discuss the study findings in relation to differences in the treatment pathways or care settings seen in the analytical sample and the Canadian healthcare system, as it may impact on the relevance of results to the Canadian context.

Section 13: Summary of Recommendations

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

Section 14: Limitations

Overview:

A limitations section should be included in the submission. However, much of this section can refer to previously mentioned limitations in the preceding sections (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 research 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 9 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 the plausibility of results by using causality frameworks, such as the Bradford Hill criteria[39] (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 healthcare data, and very small effect measures can be statistically significant without having meaningful implications for clinical practice.

Section 14: Summary of Recommendations
95. Provide a consideration of limitations of the study, including the data source, missing data, bias and confounding, imprecision/sample size limitations, and whether results are clinically meaningful 96. Discuss the plausibility of results and whether results could be due solely to bias, chance, or confounding

1141

1142

Forward-Looking Statement and Conclusion

The development of this guidance was founded on the following principles: 1) To ensure that regulators and HTA agencies have sufficient information to evaluate a submission for its appropriateness of use for decision-making; 2) To provide core reporting standards for RWE submissions that align with global standards; 3) To prioritize transparency in reporting while accounting for practical challenges related to RWD and RWE.

This guidance was developed during a time of immense change in the fields of RWD and RWE where there is ongoing advancement and development in data sources and methods in parallel to a surge in novel health technologies and medical interventions. As such, this guidance was written in a manner that allows for flexibility given ongoing changes in the landscape of RWE and RWD. This document will likely require continuous updating and extensions as the field evolves. Lastly, guidance development is the first step to establishing the use of RWE in decision-making for HTA and regulatory decisions. Future efforts can leverage these core reporting standards to provide guidance on how and when RWE can be used in HTA and regulatory decision-making.

References

1. CADTH. Real-world evidence: a primer. 2022: https://www.cadth.ca/real-world-evidence-primer. Accessed 2022 Oct 7.
2. 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. Accessed 2022 Oct 7.
3. 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.
4. Sobel RE, 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.
5. Yuan H, Ali MS, Brouwer ES, et al. ISPE comparative effectiveness research special interest group. 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.
6. 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.
7. 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.
8. Guyatt G, Rennie D, Meade MO, Cook DJ. eds. *Users' guides to the medical literature: a manual for evidence-based clinical practice*, 3rd ed. New York: 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.
9. Consort. Transparent reporting of trials. 2010: http://www.consort-statement.org/. Accessed 2022 Aug 31.
10. Prisma. Transparent reporting of systematic reviews and meta-analyses. 2020: https://www.prisma-statement.org/. Accessed 2022 Aug 31.
11. STROBE. Strengthening the reporting of observational studies in epidemiology. 2022: https://www.strobe-statement.org/. Accessed 2022 Aug 31.
12. Equator Network. Enhancing the quality and transparency of health research. [no date]: https://www.equator-network.org/. Accessed 2022 Aug 31.
13. 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*. 2021:12;372:m4856.
14. 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.

15. Bryan J, Daniel G, Silcox C, 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.
16. 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.
17. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915-20.
18. Chikuba K, Saito T, Kaketani K, et al. [The intensity of hst-1 gene amplification and malignant potential of esophageal carcinoma in relation to intratumoral heterogeneity]. *Gan To Kagaku Ryoho*. 1993;20(6):763-766.
19. 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.
20. 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.
21. 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.
22. Wasserstein RL, Lazar NA. The ASA statement on *p*-values: context, process, and purpose. *Am Stat*. 2016;70(2):129-133.
23. Mustafina RH. Enhancement of the radiation effect by fibrinolysin and quantitative pattern of the change of tumor radiosensitivity. *Neoplasma*. 1984;31(3):263-269.
24. 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.
25. Velentgas P, Dreyer NA, 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.
26. Johnston BC, Patrick DL, Devji T, et al. Chapter 18: Patient-reported outcomes. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). London: Cochrane; 2022: https://training.cochrane.org/handbook. Accessed 2022 Oct 27.
27. Tettamanti G, Ghidoni R, Trinchera M. Recent advances in ganglioside metabolism. *Indian J Biochem Biophys*. 1988;25(1-2):106-111.
28. Rothman K, Greenland S, Lash,TL. *Modern Epidemiology*, 3rd Edition. Philadelphia: Lippincott Williams & Wilkins; 2008.
29. Strom BL, Kimmel SE, Hennessey S. *Pharmacoepidemiology*, Sixth edition. New York: John Wiley & Sons Ltd; 2019.
30. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492-499.
31. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. *JAMA.* 2016;316(17):1818–1819.
32. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol*. 2012;41(2):514-520.

33. 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.
34. International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E9(R1) estimands and sensitivity analysis in clinical trials. 2019: https://database.ich.org/sites/default/files/E9%28R1%29%20Training%20Material%20-%20PDF_0.pdf. Accessed on 2022 Oct 25.
35. 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):363-9.
36. Yao XI, Wang X, Speicher PJ, Hwang ES, Cheng P, Harpole DH, Berry MF, Schrag D, Pang HH. Reporting and guidelines in propensity score analysis: a systematic review of cancer and cancer surgical studies. *J Natl Cancer Inst*. 2017;109(8):djw323.
37. Halsey LG. The region of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biol Lett Royal Society*. 2019;15.
38. 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.
39. Hill, AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58:295–300.

Appendix A: Submission 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 research 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. Reference to a priori protocol(s)		
	8. Describe all study team members including the role of patient partners		
	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		
	12. Describe important information to contextualize the data source, including:		
	a. Type of care setting b. Geographical location		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
Section 2 Setting and Context	13. Describe all relevant study period dates, including periods of recruitment, exposure, follow-up, and data collection		
	14. Report the restrictions and limitations on codes to those identified from certain settings		
	15. Clearly identify missing areas of insight in the data collection		
	16. For studies that propose the use of a data source from a country other than Canada, provide: a. A rationale for selecting the data source b. An explanation of how all these factors might affect the generalizability of the study results to the Canadian population c. Background information about the healthcare system d. A description of prescribing and utilization practices e. Information on the market availability		
Section 3 Data Specifications: Access, Cleaning Methods, and Linkage	17. Describe the extent to which the investigators had access to the database population used to create the study population		
	18. Provide information on the data-cleaning methods used in the study. If possible, share any data cleaning code leveraged. If not provided, justify		
	19. Describe the usage of data and consent for data sharing. Provide consent documents if relevant		
	20. Describe data collection methods		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
	21. Quality of the data and relevant metrics to assess the data quality should be reported		
	22. Describe any variability between data sources and the impact of changes over time in the data		
	23. Describe if any data linkage was conducted and the methods used for the linkage		
	24. Report who (organization) performed the linkage		
	25. Describe the performance characteristics of data linkage and the number of individuals linked		
Section 4 Data Sources	26. Provide and describe all data sources		
	27. Describe the characteristics of the health setting and context of data collection		
	28. Describe details of data continuity and completeness		
	29. Include the names, dates and/or version numbers of when data were extracted for research use by the data vendor or organization		
	30. 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		
	31. Provide source(s) of data for each variable of interest		
	32. Describe how variables of interest were measured and if they have been		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
	adjudicated or validated in the population of interest		
Section 5 Participants	33. Provide inclusion criteria used to identify the study population		
	34. Justify exclusion criteria and how ~~it~~ they may affect the overall interpretation of the research		
	35. Describe study population characteristics relative to the target Canadian population		
	36. Provide all codes or algorithms used to define the inclusion and exclusion criteria		
	37. Specify the time period (e.g., lookback window) over which inclusion and exclusion criteria were assessed		
	38. Recommendations for specific study designs: a. 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 b. For prospective cohort studies, describe recruitment processes c. For case-control and case-crossover studies, provide details of case and control ascertainment, the source population for nested studies, sampling methods, matching criteria		
	39. Report the numbers of participants at each stage of the study and reasons for non-participation. Consider illustrating this information using a flow diagram		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
Selection 11 Participant Characteristics	40. Provide characteristics of study participants		
	41. Indicate missing data for each variable of interest		
	42. Statistically compare treatment or exposure groups		
	43. Specify the number of participants included in each analysis and the analysis strategy (e.g. per-protocol, ITT)		
Section 6 Exposure Definitions and Comparators	44. Define the comparator group(s) (e.g., active comparator, historical comparator)		
	45. Provide justification for the comparator used including potential implications on the study findings		
	46. Discuss any changes in patterns of use of the exposure and comparator(s) over time and how it may affect the results		
	47. Specify how adaptations to the intervention and/or comparator were permitted and recorded		
Section 7 Outcomes	48. Report definitions for all study outcomes (primary, secondary, and exploratory)		
	49. Provide a rationale for the outcomes studied and discuss relevant outcomes not included in the study		
	50. Provide information about the validity of all outcome definitions		
	51. Describe whether the timing of the outcome can be accurately measured		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
	52. Specify whether the outcome studied is a surrogate measure of a clinical (patient-centered) outcome and, if so, the strength of the relationship between the surrogate outcome and major clinical outcome(s) of interest		
	53. Discuss whether outcome misclassification could occur between treatment groups		
	54. Report whether a control outcome(s) was used and justify the control outcome(s) selected		
Section 8 Variables and Data Dictionary	55. Provide a data dictionary that includes information on data sources, validity, and definitions for all variables, as applicable		
	56. Specify definitions and lookback windows for all variables		
	57. Report whether any variables could be time-varying (e.g., how the variable could change over time and when it was re-defined in relation to time-varying exposures)		
	58. Report important variables that could not be captured and their anticipated impact on study results		
	59. Provide information on deviations from the *a priori* protocol		
	60. Report all procedures used to address potential sources of bias		
	61. Specify how potential sources of bias could influence the outcomes of the analyses		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
Section 9 Bias, Confounding, and Effect Modifiers/Subgroups Effects	62. Specify variables that were considered known or potential confounders in the analysis		
	63, Describe how confounder variables were selected and if they were informed by a causal diagram		
	64. Describe and compare the distribution of measured baseline confounding variables between treatment groups		
	65. Report whether any potential confounders could not be measured and specify the anticipated impact of these confounders on study results		
	66. Report whether time-varying confounding was considered and if not, justify why not		
	67. 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		
	68. Specify known or potential effect modifiers		
	69. Describe any effect modification or subgroup analyses that were conducted and if they were specified *a priori*. If effect modification/subgroup analyses by main demographic variables (age, sex, race/ethnicity) or by pertinent comorbidities were not employed, justify why these analyses were not needed		
	70. If effect modification/subgroup analyses by main demographic variables (age, sex, race/ethnicity) or by pertinent comorbidities were employed,		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
	describe the methods used and present separate results for each subgroup		
Section 10 Statistical Methods	71. Indicate the software(s) used for the statistical analysis including software package, version, and analytic tools employed (e.g., macros)		
	72. Provide access to the statistical code used or if the code cannot be shared, explain why not		
	73. Report all statistical methods used and justify their selection, including, as applicable:		
	a. All variables included in regression models		
	b. The method of variable selection for regression models		
	c. Methods used to control for confounding		
	d. Methods used to account for missing data		
	e. How follow-up time and changes in exposures were handled		
	f. Subgroup analyses and effect modification		
	g. As applicable: stratification, propensity score estimation and assumptions, meta-analysis methods, validity of instrumental variables		
	74. Quantify the precision of all estimates using confidence intervals		
	75. Report the threshold of statistical significance used		
	76. Summarize key results (estimated effect measures, measures of precision) with reference to each study objective and/or hypothesis for primary		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
Section 12 Study Findings	and secondary outcomes and delineate these results by each treatment or exposure group		
	77. Provide numbers of outcome events or summary measures of outcomes (or exposures in case-control studies)		
	78. Report both absolute and relative effect measures for binary outcomes including their measure of precision		
	79. Report category boundaries when continuous variables are categorized and consider translating estimates of relative risk into absolute risk		
	80. Report unadjusted and adjusted estimates including their measure of precision and confounders used for adjustment		
	81. Report other pre-specified analyses conducted (e.g. subgroup analyses, interactions, sensitivity analyses)		
	82. Describe any unplanned analyses performed secondarily (not defined *a priori*) and indicate these as exploratory		
	83. Avoid selective reporting of results		
Section 13 Interpretation and Generalizability	84. Provide an interpretation of the primary and secondary study results, as applicable		
	85. Interpret the findings from adjusted and unadjusted results, as applicable		
	86. Discuss the precision of the effect measure(s)		
	87. Discuss how potential biases and sensitivity of study assumptions may		

Section	Checklist Item	Reported on page number(s)	If not reported or not applicable, justify why
	impact the results and subsequent interpretation		
	88. Discuss the implication of findings for clinical practice, including the risk-benefit profile of the treatment, if applicable		
	89. Interpret study findings in relation to current literature		
	90. Discuss the generalizability (external validity) of study results to the Canadian population		
Section 14 Limitations	91. Provide a consideration of limitations of the study, including the data source, missing data, bias and confounding, imprecision/sample size limitations, and whether results are clinically meaningful		
	92. Discuss the plausibility of results and whether results could be due solely to bias, chance, or confounding		

Appendix B: Methods

METHODS

A two-phase process to develop this guidance document was used. **Phase 1** aimed to conduct an environmental scan of international RWE evidence and an evidence mapping process in order to develop candidate items to be included in the submission guidance. **Phase 2** used a modified Delphi process to select final recommendations from the list of candidate recommendations, include additional relevant items if required, and provide guidance for implementation of recommendations and special considerations for the Canadian context, as appropriate.

Phase 1, Part 1: Identification of Documents on RWE (Environmental Scan)

Articles were identified in the 2020 Environmental Scan by CADTH: "Use of Real-World Evidence in Single-Drug Assessments" that aimed to identify Canadian and international agency guidance, reporting tools, and policy statements related to RWD and RWE.[7] Detailed methods used in this environmental scan are described in the original document. Briefly, 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.

This environmental scan was extended to identify potential additional documents by: 1) using a citation-search method of documents included in the environmental scan; and 2) consulting the expert panel to identify articles that had not yet been included or were currently in development. In total, we identified 37 documents for review (**Appendix C**).

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). For each category, we developed a matrix with sub-categories (e.g., protocol, exposures) to which recommendations were mapped/organized. Additional matrix categories were added if a new theme was identified within a document. 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.

Appendix Table 1. Categories of Data Extraction Matrices.

Component	Categories to which recommendations were mapped
Reporting of RWE	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 RWE	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 panel and stakeholders. The other two investigators (KH, TA) then independently reviewed the duplicate removal and standardization of language to ensure that items were 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 14 major themes that correspond to the different sections in the guidance document (e.g., Participants; Bias and Confounding).

Part 2: Delphi/Consensus Process

Authorship Team: The Delphi process was led by three 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 grouped into 14 themes were programmed into an online questionnaire. Canadian and international RWE experts and stakeholder participants were asked to determine the importance of including each item into the submission guidance document. Each item was ranked on an anchored scale of 1 to 4 where 1 indicated "*not important*" and 4 *"indicated 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 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), their scores, 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 an agreement of ≥70% of respondents to include were included in the guidance document. The same level of agreement was used for items to exclude. Items that generated <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) where each item that did not generate consensus was discussed. For each item, participants voted on whether to "*include*", "*omit*", or "*revise*" each item using an online polling feature. Items that generated ≥70% 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 (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 and the authorship team took detailed notes.

Following the first discussion, the authorship team drafted the first iteration of the guidance document. Items confirmed by the expert panel were a) inputted into a reporting checklist and b) elaborated upon to provide additional guidance on how to implement the item (i.e., details to guide the submission). The guidance document was circulated to the expert panel for review; The initial review was 2 weeks. Participants were invited to provide feedback on the document via email or on the shared document. Additional feedback was collected from internal stakeholders, specifically 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/which items should be revised was held virtually on September 20th, 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 this version. Next steps will include public posting of the document and collection of written submissions from all stakeholders. These documents will be compiled and reviewed by the internal team for feasibility of change to the document. Suggestions that are out of scope or will be excluded. Feedback will be presented to the external panel in January 2023 for input on inclusion or changes to the guidance into the final draft.

Participants: The expert panel was purposefully selected to include 10 Canadian members 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 a 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 guidances. All experts had to declare potential conflicts of interests and align with CADTH conflicts of interest policy.

Results: A total of 13 respondents completed the survey. See **Appendix Table 2** for results. Sections that generated 100% agreement included: *participants, exposure definitions and comparators, effect modifiers* and *study findings*. Other sections generated reasonable to high levels of agreement (67-92%). Sections 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 observers). Thirty items were discussed during the June 22 meeting and 6 items were discussed asynchronously after the call due to a lack of time.

A total of 14 participants from the expert panel and 13 internal stakeholders (CADTH and Health Canada) reviewed the first iteration of the RWE guidance document. Fifteen discussion items were put forward during the September 20th 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.

Appendix Table 2. Results from first expert consensus survey.

Section (*n=13 out of 15-panel members*)	Number of questions	Overall Agreement	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
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**

Appendix C: Documents Reviewed for Candidate Recommendations on RWE/RWD Reporting for Expert Survey.

1. FDA - guidance for RWE*
2. FDA - Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data*
3. INESS*
4. EMA Guidance for Registry-Based Studies (2020)*
5. EnCEPP Pharmepi Guide - v9 2021*
6. EMA Guidance for the format and content of the final study report of non-interventional post authorisation safety studies*
7. AHRQ Patient registry Guide - 4th edition*
8. REQuest Tool (2019)*
9. CONSORT*
10. STROBE*
11. RECORD*
12. RECORD-PE*
13. ISPOR Guideline 2020 - "Improving transparency to build trust..."*
14. Wang 2017 - Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0 *
15. HTAI
16. NICE guidelines
17. Elements of Real World Data/Evidence Quality throughout the Prescription Drug Product Life Cycle (2019)*
18. eunethta Guideline: COMPARATORS & COMPARISONS (2015) *
19. CADTH 2020*
20. ARROWs 2021*
21. CASP cohort studies*
22. CASP case control studies*
23. newcastle ottawa*
24. MINORS (2003)*
25. Liaw 2021: Quality assessment of real-world data repositories across the data life cycle: A literature review
26. HIDQF (Kahn 2016)*
27. Duke-Margolis Center for Health Policy, 2018*
28. DUKe- RWD 2018*
29. Miksad and Abernethy*
30. Berger 2017 ISPOR-ISPE*
31. ISPE GPP 2015*
32. PCORI Methodology Standards Checklist*
33. GRACE checklist/GRACE Principles*
34. JBI-MAStARI*
35. ROBINS-I tool*
36. START RWE*
37. Bolislis 2020

*denotes that candidate recommendations were extracted from the document.