

## General Comments

ISPOR, the Professional Society for Health Economics and Outcomes Research, appreciates the opportunity to provide comments on CADTH’s consultation on its Real-World Evidence (RWE) Reporting Guidance. Our response is based on comments from several of ISPOR’s member groups having expertise in this area.

Overall, the document is a valuable summary of the best practices and methodological considerations for RWE research and reporting. It will help to support the appraisal of RWE for regulatory and HTA purposes in Canada and beyond. The guidance presentation is well organized, any potential limitations to reporting certain details have alternative options presented, and the provided reporting checklist is extremely useful.

The guidance could be improved by providing further example diagrams discussed throughout for users to adapt and by providing external links to relevant case studies or submissions as specific examples of high-quality research and a successful submission. Including additional methodological considerations, case studies related to economic analyses, and the evidence required for HTA would be helpful. We have provided specific details on these suggested improvements in the relevant sections below.

### Section/line-specific comments

Line 63: We recommend changing “medical technologies” to “health technologies” for consistency throughout the document.

Lines 96 - 98: We recommend including additional uses of RWD within the HTA context, such as a) serve as input parameters in economic modeling, including the cost of treatments, and b) gain insights into current prescribing trends to determine treatment mix in budget impact assessment.

Lines 100-101: We recommend adding “patients from diverse ethnic groups” as another advantage.

Lines 155-156: Correction “regulatory... decision-making.”

Lines 234-242. This section states that the aim and research question must be clearly reported. We suggest that it also states that any hypothesis that will be tested in the study be reported.

Line 254. It is recommended for a study design diagram be included in this section to improve transparency and interpretability. Including template study design diagrams for various RWE study designs or checklists that identify what aspects of the study design this diagram should include would be a useful resource to include in a future version of this reporting guidance.

Lines 254-259. The follow-up window, including the details of how follow-up ends (e.g., the first event of the outcome or death), may be included in the diagram. This diagram can be accompanied by relevant footnotes, which can provide further details about what is defined in each of the assessment windows.

Lines 261-263: The development of a prior protocol and statistical plan is a critical part of any reproducible research. We suggest including the development of a priori protocol as a requirement and not a recommendation. For transparency and consistency across protocols, a template such as HARPER could be suggested (<https://pubmed.ncbi.nlm.nih.gov/36215113/>). In addition, please consider making suggestions for the registration platform of the a priori protocols, with attestation to its a priori nature preferred.

Lines 304-305: The study should mention the specific version of the database and the date of the last update.

Line number 311-312: Besides mentioning missing data components, the study should also mention any mechanism for their missingness (to help justify their subsequent treatment as missing completely at random, missing at random, or missing not at random) and the extent of missingness in the dataset to provide information about the completeness of the dataset.

Lines 334-402. It was good to see the recommendation that all code or the references to all external code used should be included, but it may be difficult to provide all such code for multiple reasons. Further discussion on the alternative options when it is not possible/feasible to provide code could be further detailed. For example, one alternative may be to detail which packages or functions within the code were used throughout the data cleaning and analysis stages. This is something that the community can benefit from- to explore additional RWE analysis methodologies, and those who review the submission could still assess the suitability of the methodology used in the submission. The feasibility of other recommendations, including audit trails, could also be considered for different types of research investigators.

Lines 340-341. Suggest adding at least a recommendation that a data specification document should be created and available with the submission of study information; the development of the data specification(s) should be clearly documented with versions and dates of the respective documents.

Lines 360-367. The data cleaning process should also report the identification of any duplicates and the process to evaluate any data collection errors, including entry, measurement, integration, and summarization errors. Investigators should report their strategies to minimize missing data. If patient experience data were involved, the study should report summary statistics of missing data frequencies and percentages, stratified by important subgroups.

Line 393: If primary data collection is involved, it should mention any testing of the reliability/validity of the data collection instrument, with supporting documentation, as well as limitations to its use. In addition, considerations of re-identification risk analysis when managing pseudonymized data that has to be sent for analysis could be mentioned.

Line 400. Point 20 of the summary – does “describe the data collection methods” apply to primary data collection only or all types of data collection?

Lines 407-418. The data source description should also address the validity and reliability of the dataset, particularly if the dataset is new or not widely used by research groups. If a novel data source was used, the study might also mention any comparison of estimates obtained from the database with estimates obtained from similar but well-known or common datasets to indicate any overestimation or underestimation inherent in the dataset.

Lines 490 – 494: It was noted that all relevant code lists should be reported, and validation methods should be cited or detailed, but we believe this should be further emphasized, considering how much code lists can impact the final results of RWE studies. A clear stance on how important validation studies of code lists and reporting code lists are for CADTH would be particularly helpful. Depending on the intended audience, it may also be helpful to define code lists as well as target populations more extensively than currently provided. In addition, it would be useful to include a glossary defining technical terms used throughout the guidance, such as “code lists” and “target population,” as an appendix or at the forefront of the document in order for those with less technical experience to have these definitions readily available and establish a uniform understanding as to what is meant by these terms.

Lines 527-528. Since the under-inclusion of a group cannot always be adjusted for, we suggest changing the second sentence to “Acknowledge whether some patient groups may not be *well*-represented and, if so, how *under*-inclusion of these groups may *affect* the external generalizability of research findings.”

Lines 536-539: We suggest adding recommendations related to sex/gender for the research conduct and reporting, in line with changing requirements/best practices in Canada, as accepted in national guidelines.

Lines 565-567: Matching the cohorts on the observed variables is critical to comparative effectiveness. However, multiple statistical tests on baseline characteristics are not informative and may be misleading for the judgment of goodness of the matching procedure. We suggest replacing “statistical comparison” with “comparison and evaluation of participant’s characteristics between treatment and exposure groups...”.

Lines 606-639: Dose modification should be more specifically called out under Exposures (not only time of exposure but also the extent).

Lines 606-672: There should be a section outlining biomarker use and assays, given that for several therapies, treatment depends on the expression of a specific biomarker, which may be assessed through different commercially available assays.

Lines 632-635: For exposure, we recommend specifying the procedure and parameters used to define the start and the end dates/period.

Lines 724-725. If an outcome is self- or observer-reported, specify whether a validated instrument was used and reference the validation studies and respective details (incl. version, administration mode, linguistic validation, etc.). If a validated instrument was not used to capture the outcome, the rationale for selection, administration, and details on the instrument’s scoring should be described.

Lines 744 & 1068, on causality. While not drawing “conclusions” about causality is technically correct, it should also be recognized that some level of causal inference is generally the point of comparative effectiveness studies with RWD. How can they inform decision-making without any level of causal inference from the results?

Lines 822-823. Suggest modifying this sentence to “If multiple data sources, *or multiple years of the same data source*, were used, report any differences ... between sources, *and how those differences were reconciled or addressed*.”

Lines. 987-991. The first and fourth sentences of this paragraph do not seem consistent.

### **Appendix A, B & C (Lines 1322-1540)**

A key highlight of this guidance was the inclusion of the reporting checklist for each section summary in Appendix A. One additional resource that was included in the NICE RWE framework (June 2022) that would prove useful in this guidance would be integrating example methodology reporting text from case studies of different RWE study designs to indicate the level of detail expected as well as some existing RWD sets accepted. CADTH should consider creating those links in the Summary of Recommendations at the end of each session and constitute the content of the final checklist in the form of a table available in APPENDIX A. However, the numbering & name of some sections in the text (Summary of Recommendations) and appendix A seem not aligned. The Summary of Recommendation from section 6 seems incorrect, and this error propagates until the end of the text.