PHILIPPINE HTA GUIDANCE DOCUMENT

FOR THE USE OF REAL-WORLD EVIDENCE IN CLINICAL ASSESSMENTS OF HEALTH TECHNOLOGIES

OCTOBER 2024

Department of Science and Technology (DOST)
Health Technology Asessement (HTA) Division

Philippine Guidance Document on the Use of Real-World Evidence for Clinical Assessment of Health Technologies

University of the Philippines National Institutes of Health - Institute of Clinical Epidemiology
in collaboration with
Health Technology Assessment Division, Department of Science and Technology, Philippines
October 2024

Acknowledgements

Technical Editors
Carol Stephanie C. Tan-Lim, MD, MScCE, DPPS, DPSAAI
Marissa M. Alejandria, MD, MSc, FPCP, FPSMID

Contributors
Natasha Ann R. Esteban-Ipac, MD, GDip (Clin Epi), FPPS, DPSAMS
Ian Theodore Cabaluna, RPh, MD, GDip (Clin Epi), MSc (Clin Epi)
Anna Angelica Macalalad-Josue, MD, FPCP, FPCEDM, MSc (cand)
Howell Henrian G. Bayona, MSc, RSLP
Aldrich Ivan Lois D. Burog, MD, MSc (cand)
Kim L. Cochon, PhD
Myzelle Anne J. Infantado-Alejandro, MSc (cand)

Copy Editor
Larraine Franchesca L. Lopez, OTRP, MD

Project Manager
Mark Dale S. Imbag, MD

External Reviewers
Wanrudee Isaranuwatchai, PhD
Lucylynn Lizarondo, PhD
Wee Hwee-Lin, PhD

Pilot Assessors
Kerwyn Jim C. Chan, RSLP, MSc
Kimberly Mae C. Ong, MD, MSc, FPSOHNS
Zaira Nina D. Duque, RN MD DPPS
Frangelo Conrad Tampus, MD
Cary Amiel G. Villanueva MD MPH (Glas)

Pretesters of External Reviewers and Pilot Assessors' Questions
Maria Vanessa Sulit RN, MSc (Clin Epi)
Adovich Rivera, MD, PhD

Conflict of interest declaration: All individuals involved in the development of the methods guide have no conflicts of interest to declare related to this activity.

Contact information
University of the Philippines National Institutes of Health
upm-nih@up.edu.ph

Health Technology Assessment Division – Department of Science & Technology
hta@dost.gov.ph
htanominations@dost.gov.ph
htaresearch@dost.gov.ph

To cite this document, we recommend citing the journal article that describes the development of this document: Tan-Lim CC, Cabaluna IG, Infantado-Alejandro MJ, Imbag MS, Lopez LI, Alejandria MM. Development of the Philippine Guidance Document for the Use of Real-World Evidence for Clinical Assessment of Health Technologies. [Submitted for publication]

Table of Contents

Abbreviations

aOR Adjusted odds ratio
APA American Psychological Association
aRR Adjusted relative risk
ArRoWs Assessment of Real-World Observational Studies
CADTH Canadian Agency for Drugs and Technologies in Health
CDA Canada's Drug Agency
CENTRAL Cochrane Central Register of Controlled Trials
CONSORT Consolidated Standards of Reporting Trials
EHR electronic health records
EMA European Medicines Agency
EMR Electronic medical records
GIV Generic Inverse Variance
GRADE Grading of Recommendations Assessment, Development, and Evaluation
HMIC Health Management Information Consortium
HTA Health Technology Assessment
ICC Intraclass Correlation Coefficient
ISPOR International Society for Pharmacoeconomics and Outcomes Research
KII Key Informant Interviews
MD mean difference
MeSH Medical Subject Headings
MINORS Methodological index for non-randomized studies
NHS EED National Health Service Economic Evaluation Database
NHS PROMs National Health Service Digital Patient-Reported Outcomes Measurement
NICE National Institute for Health and Excellence
NIH National Institute for Health
NIHR National Institute for Health Research
NMA Network meta-analysis
PCT Pragmatic Clinical Trials
PICO Population, Intervention/Exposure, Comparator, Outcome
PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROMIS Patient-Reported Outcomes Measurement Information System
PROTECT Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium
RCT Randomized Controlled Trials
RECORD Reporting of studies Conducted using Observational Routinely-collected Data
RTI Item Bank Research Triangle Institute Item Bank
ROBINS-E Risk of Bias in Non-randomized studies- of Exposures effect
ROBINS-I Risk Of Bias In Non-randomised Studies - of Interventions
RWD Real World Data
RWE Real World Evidence
SMD Standardized Mean Differences
STROBE Strengthening the Reporting of Observational Studies in Epidemiology
UHC Universal Health Care
US FDA United States Food and Drug Administration
US NTIS United States National Technical Information Service
WHO World Health Organization

Chapter I

Introduction

1.1 What is the purpose of this document?

1.2 What is the scope of this document?

1.3 Who is the target audience of this document?

1.4 What is the process of development of this document?

1.5 How will this document be updated?

Chapter Summary

This document provides methodological standards for the use of real-world evidence in the clinical evaluation of health technologies in the Philippines, including searching, appraising, and utilizing RWE. The primary target audience of this document are researchers aiming to produce HTA reports for the Department of Health and PhilHealth. There were two phases in the development of this document. Phase 1 was a comprehensive, systematic review of relevant literature. Phase 2 was a validation study through expert consultation using key informant interviews, and pilot assessment of the methods guide. Continual evaluations will be done to determine the need for an update of this document, depending on the developments in the research and healthcare landscape.

1.1 What is the purpose of this document?

Currently, there are a variety of innovations to improve healthcare and healthcare delivery systems. Such health technology encompasses medicines, vaccines, devices, procedures, and systems. Under the Universal Health Care (UHC) Act, all health technologies should undergo health technology assessment (HTA) before implementation to ensure the responsible use of health technologies financed by the government.[1] Health technology assessment (HTA) is defined in the Philippine HTA Methods Guide as a systematic evaluation of health-related technologies using a multidisciplinary approach incorporating the clinical, economic, organizational, social and ethical evaluation of a health technology.[2]

The current Philippine HTA Methods Guide provides broad guidelines in the conduct of systematic reviews and economic evaluations to produce standard HTA reports for health technologies. However, certain health technologies such as orphan drugs for rare diseases, emerging health technologies, hospital medical equipment and devices, diagnostic and screening tools, and preventive and promotive health services may require different types of evidence as well as different methodological standards to assess their potential value to our health care system.

With the increased availability of real-world data (RWD), the use of real-world evidence (RWE) to inform healthcare decision-making has been gaining increased salience.[3] Although acceptance of RWE in regulatory decision-making is not yet universal, several countries have integrated RWE in their HTA process. [4] Canada, the United Kingdom, the United States, Europe, and some Asian countries have developed guidance documents on the use of real-world evidence in HTA to ensure the appropriate use and application of RWE in the HTA process. [5-10] This document aims to develop methodological standards for the use of RWE in the clinical evaluation of health technologies in the context of the Philippines. This specialized methods guide shall be interpreted and used together with the overarching HTA methods guide - the 2nd edition of the Philippine HTA Methods Guide. [11]

1.2 What is the scope of this document?

This document provides guidance for searching, appraising, and utilizing RWE in conducting clinical evaluation of health technologies as part of HTA in the Philippines. This methods guide does not cover other important aspects of HTA, including economic, organizational, social, and ethical evaluation which are available in the 2nd edition of the Philippine HTA Methods Guide.

1.3 Who is the target audience of this document?

The primary target audience of this document are researchers (including the assessors HTA Division, or commissioned researchers forming the External Assessment Groups) aiming to produce HTA reports which are used as evidentiary bases for the HTA Council in developing coverage recommendations for the Department of Health and PhilHealth. This methods guide is particularly relevant for researchers aiming to perform clinical evaluation of health technologies where randomized controlled trials may not be feasible, sufficient, or relevant to provide a comprehensive assessment of the potential value of the health technology to the Philippine healthcare system.

This document also serves to guide the HTA Council and other decision-makers on the methodological standards in using RWE for HTA to ensure the validity and reliability of using RWE in informing decision-making. Although HTA reports are produced internally or by commissioned external assessment groups, industries, and other stakeholders may also use this guide to provide submissions that adhere to the methodological standards prescribed in this document.

This document may also serve as a useful reference for other stakeholders affected by the implementation and monitoring of health technologies, including health authorities, healthcare providers, healthcare organizations, and patient organizations.

1.4 What is the process of development of this document?

There were two phases in the development of this document. Phase 1 was a comprehensive, systematic review of the available HTA methods guide of other countries as well as literature related to the use of RWE in the clinical evaluation of health technologies. A systematic search of electronic medical databases and HTA organization websites, and a manual search of references were done. Content experts were also contacted to access unpublished research. Based on the results of the systematic review, a draft HTA methods guide on the use of RWE was created.

Phase 2 was a validation study through: (1) expert consultation using key informant interviews (KIIs); and (2) a pilot assessment using the methods guide (Figure 1). Experts were selected as key informants if they had previous experience in creating a methods guide on the use of RWE for HTA or regulatory decision-making, or if they had expertise in using RWE for the clinical evaluation of health technologies. Three experts were consulted. Pilot assessors were chosen among the potential users of the methods guide. Five pilot assessors with varying experience in preparing evidence summaries for the clinical evaluation of health technologies were selected.

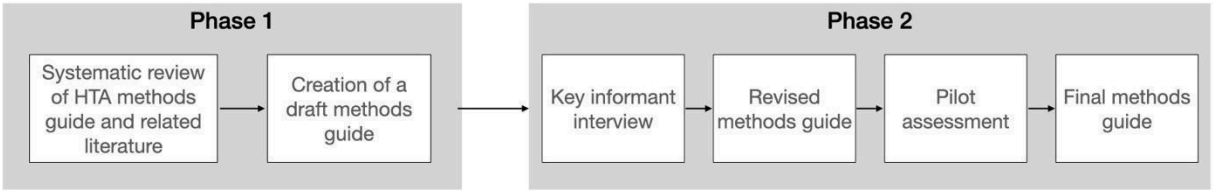

Figure 1. Process of development of the methods guide

1.5 How will this document be updated?

This is the first edition of the HTA methods guide on the use of RWE for clinical evaluation of health technologies in the Philippines. Due to the evolving nature of RWE, this document will be periodically updated to reflect advancements in RWE methodology and RWD sources, users' feedback, and the changes in the Philippine healthcare setting. Users of this methods guide are encouraged to provide feedback through the Health Technology Assessment Division – Department of Science & Technology (hta@dost.gov.ph)

Similar to international HTA methods guides and the Philippine HTA Methods Guide, there is no prescribed time interval for the update to be performed. Continual evaluations will be done depending on the developments in the research and healthcare landscape.[5-7]

References:

1. Congress of the Philippines. Republic Act No. 11223. An act instituting universal health care for all Filipinos, prescribing reforms in the health care system, and appropriating funds therefor. Official Gazette. 2019 Feb 20 [accessed 2024 Jun 20] https://www.officialgazette.gov.ph/2019/02/20/republic-act-no-11223/
2. Health Technology Assessment Unit, Department of Health Philippine HTA Methods Guide. 1st ed. 2020
3. Kamusheva M, Németh B, Zemplényi A, Kaló Z, Elvidge J, Dimitrova M, et al. Using real-world evidence in healthcare from Western to Central and Eastern Europe: a review of existing barriers. J Comp Eff Res. 2022 Aug;11(12):905-913. doi: 10.2217/cer-2022-0065. PMID: 35726611.
4. Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: The changing landscape of the regulatory approval process. Pharmacoepidemiol Drug Saf. 2020 Oct;29(10):1201-1212. doi: 10.1002/pds.4932. PMID: 31823482; PMCID: PMC7687110.
5. National Institute for Health and Care Excellence. NICE real-world evidence framework [Internet]. [cited 2024 Jun 20]. Available from: https://www.nice.org.uk/corporate/ecd9/chapter/overview
6. Canada's Drug and Health Technology Agency. Guidance for Reporting Real-World Evidence [Internet]. [cited 2024 Jun 20]. Available from: https://www.cadth.ca/guidance-reporting-real-world-evidence
7. US Food and Drug Administration. Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products [Internet]. [cited 2024 Jun 20]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug
8. Institut national d'excellence en santé et en services sociaux (INESSS). Integration of Real-World Data and Evidence to Support Decision-Making in the Pharmaceutical Sector. 2022. https://www.inesss.qc.ca/en/publications/publications/publication/integration-of-real-world-data-and-evidence-into-assessments-to-support-decision-making-in-the-pharmaceutical-sector.html.
9. Haute Autorité de santé. Real-world studies for the assessment of medicinal products and medical devices. 2021. https://www.has-sante.fr/jcms/p_3284524/en/real-world-studies-for-the-assessment-of-medicinal-products-and-medical-devices.
10. Lin L, Ahn J, Bayani D, Chaiyakul M, Lee Y, Tan Y, et al. Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in ASIA. 2020. https://hiper.nus.edu.sg/realise-guidance/. Accessed February 17, 2024.
11. Department of Science and Technology - Health Technology Assessment Council (2024). Philippine Methods Guide 2nd edition. [Draft submitted for approval]

Chapter II

What is the definition of Real-World Data and Real-World Evidence?

2.1 What is real-world data?

2.2 What are the sources of RWD?

- 2.2.1 What are Primary Source RWD?
- 2.2.2 What are Secondary Source RWD?

2.3 Which RWD sources should be used for clinical evaluation of health technologies?

2.4 What is real world evidence?

2.5 What are important considerations in using RWE for clinical evaluation of health technologies?

- 2.5.1 What are the main components of RWE?
- 2.5.2 How do we evaluate the acceptability of RWE?

Chapter Summary

Real-world data (RWD) is any healthcare-related data collected from a variety of sources during routine delivery of healthcare, typically in a non-experimental setting. The sources of RWD can be classified as primary source data, where it is collected intentionally for research; or secondary source data, where it was initially obtained for another purpose and subsequently used and analyzed for research. RWD for the clinical evaluation of health technologies may be obtained from primary source data, registries and databases, or clinical records. RWE is the evidence on the utilization and potential benefits or risks of an intervention derived through analysis and interpretation of RWD using best practice methods. RWE used for regulatory decision-making should have a clear research question answered by credible RWD, with the application of appropriate analytical techniques to account for the methodologic limitations of observational studies.

2.1 What is real-world data?

The use of the terms 'Real-World Data' (RWD) and 'Real-World Evidence' (RWE) has dramatically increased in recent years; however, there is still no consensus on their standard definitions [1]. Regulatory bodies such as the United States Food and Drug Administration (US FDA) [2], European Medicines Agency (EMA) [3] and International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [4] use different, although closely related, definitions. Phrases such as "routinely collected data," "data collected during routine delivery of healthcare," and "nonrandomized controlled trial" are the common linking terms in most of these definitions [5].

For this guide, **RWD is defined as any healthcare-related data collected from a variety of sources during routine delivery of healthcare, typically in a non-experimental setting.** The exceptions are pragmatic clinical trials which are conducted in an experimental setting, but are considered sources of RWD. RWD would reflect daily activities relevant to health, including but not limited to, patients' health status, healthcare delivery, treatment utilization and health outcomes. In contrast to RWD, data from randomized controlled trials [RCTs] are obtained in an experimental setting with specific inclusion and exclusion criteria for study participants and random allocation of participants to receive the researcher-defined intervention or control.

Real-World Data (RWD) is any healthcare-related data collected from a variety of sources during routine delivery of healthcare, typically in a non-experimental setting.

2.2 What are the sources of RWD?

RWD includes qualitative or quantitative patient information, including medical history, demographic data, clinical outcomes, laboratory and imaging results, resource use and costs, and health behaviors and experiences. The sources of RWD can be classified as (i) *primary source data*, where it is collected intentionally for research; or (ii) *secondary source*

data, where it was initially obtained for another purpose and subsequently used and analyzed for research.[6,7]

2.2.1 What are Primary Source RWD?

Primary source data are purposefully collected for use and analysis for specific research studies, following well-defined scientific methods and protocols.[8] Primary sources of RWD are collected in observational research studies including prospective patient registries, prospective longitudinal cohort studies, case-control studies, and cross-sectional studies; pragmatic clinical trials; qualitative studies, and hybrid study designs (Table 1).

Table 1. Study designs for primary source RWD

Study Designs	Description
1. Observational studies	
a. Prospective patient registries	● Data is collected through registries, an organized system of obtaining clinical data and other relevant data. The data is used to evaluate clinical outcomes of specific diseases, specific health conditions, and exposed populations. ● Registries are usually categorized based on the characteristics of the population. It is usually based on a medical product, disease, or health service.[9-11]
b. Prospective longitudinal cohort studies	● Data is obtained by following up participants who have a particular condition or received a particular treatment over a period of time, and compared with a group who does not have the condition or received the treatment.[12]
c. Case-control studies	● Data is collected among "cases", or those who have the outcome of interest, and compared to "controls", or those who do not have the outcome of interest.[12]
d. Cross-sectional studies	● Data is obtained from a particular population at a single point in time or time period. This includes data derived from health surveys of patient and/or caregivers.[6, 11, 12]
2. Pragmatic clinical trials	● Pragmatic clinical trials are used to evaluate how treatments affect the outcomes in the real-world setting clinical practice. ● The design in pragmatic clinical trials closely resemble how patients are treated in clinical practice, with broader eligibility criteria, more flexible delivery of intervention, and more varied follow-up compared to a randomized control trial.[11, 13]
3. Qualitative studies	● Qualitative studies are used to obtain patient and caregivers' experiences and attitudes.[14] ● Common data collection methods used in qualitative studies include interviews, focus group discussions and observations.
4. Hybrid studies	● Data is obtained using a combination of different study designs, such as conducting a cross-sectional survey and using data from a database.[6]

2.2.2 What are Secondary Source RWD?

Secondary source data are collected for a purpose other than the intended research study; thus, there is variable quality of data. Secondary source RWD can further be divided into three: registries, clinical records and unsupervised sources. [8]

1. Registries and databases

Data collected using scientific methods and following defined protocols are classified as 'registries' or 'databases'. This includes retrospective registries (e.g. patient, disease, product or drug registries) and databases (surveillance or adverse events databases). Unlike primary source RWD, these data are not purposefully obtained for specific research studies.

2. Clinical Records

Data collected during routine medical care, not following any study protocols, but under the supervision of healthcare professionals are classified as 'clinical records'. This includes electronic medical records (EMR), also known as electronic health records (EHR) or hospital information system, and administrative records.

Data that can be collected in EMR include case notes, results of laboratory tests and medical imaging, prescription data, molecular and genomic data (usually from biobanks), and mortality data. Data that can be collected in Administrative Records include pharmacy records, health insurance, claims and billing datasets.

3. Unsupervised Sources

Data collected without the supervision of any trained healthcare professional and not following any protocol is classified as 'unsupervised sources'. This includes patient derived data from personal devices such as wearables and biosensors, smartphone health applications, and social media.

2.3 Which RWD sources should be used for clinical evaluation of health technologies?

RWD for the clinical evaluation of health technologies may be obtained from primary source data, registries and databases, or clinical records. Unsupervised sources are not recommended due to their inherent limitations, particularly the great uncertainty of the quality of this data source.[8] Since this guide focuses on clinical evaluation, qualitative studies were not included; however, qualitative studies would provide critical information for other aspects in HTA. (Table 2)

Table 2: Real World Data Sources for HTA reports[5,8,9,15]

Primary Source Data	Secondary Source Data	
	Registries and Databases	**Clinical Records**

● Observational studies
  o Prospective patient registries
  o Prospective longitudinal cohort studies
  o Cross-sectional studies
● Pragmatic clinical trials
● Hybrid studies
	● Retrospective registries
	  o Patient registries
	  o Disease registries
	  o Product registries
	  o Drug registries
	● Surveillance or Adverse Events databases
		● Electronic medical records
		  o Laboratory tests
		  o Medical imaging
		  o Prescription data
		  o Case notes
		  o Molecular and Genomic data
		  o Mortality data
		● Administrative records
		  o Pharmacy records
		  o Health insurance
		  o Claims and Billing Datasets

2.4 What is real-world evidence?

RWE refers to evidence on the utilization, potential benefits, or risks of an intervention derived through analysis and interpretation of RWD using best practice methods.[2, 5, 16] Best practice methods refer to the appropriateness and methodologic soundness of the study design and data analyses used (see section 2.5.1 for further details).

Data in itself does not equate to evidence. RWD needs to be utilized to answer a research question, with application of the appropriate study design and analytic methods, before it can be considered RWE. RWE can only be produced if the RWD used is applicable and appropriate.[5,15] This is important to note, especially in the clinical evaluation of health technologies for regulatory decision-making. The evaluation of the suitability and appropriateness of RWD and RWE for HTA use may vary depending on the specific research question or regulatory decision-making purpose.[9]

Real-World Evidence is the evidence on the utilization and potential benefits or risks of an intervention derived through analysis and interpretation of RWD using best practice methods.

2.5 What are important considerations in using RWE for clinical evaluation of health technologies?

Several factors should be considered when using RWE, including the design, conduct, monitoring, quality, ethics, and reporting quality.[11, 17] Figure 1 shows the critical components of RWE. RWE used for regulatory decision-making should have a clear research question answered by credible RWD, with the application of appropriate analytical techniques to account for the methodologic limitations of observational studies, particularly the effect of confounders.[1, 6, 16]

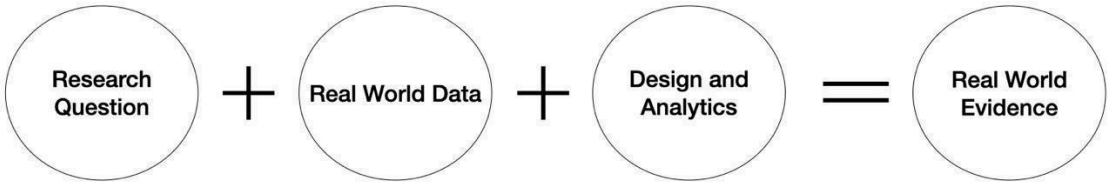

Figure 2: Components of Regulatory Grade Real World Evidence

(adapted from Delusingan[18])

2.5.1 What are the main components of RWE?

1. ***Research Question***
   The research question to be answered should be clearly defined [1], with objectives that are specific, measurable, achievable, relevant, and time-based (SMART). [10] The question should include key components such as the population, intervention/ exposure, comparator, outcome, and the time period.[1, 16]

2. ***Real World Data***
   It is crucial to evaluate the quality of RWD used in the RWE. The characteristics of high-quality RWD are: [7, 8, 16, 19]
   a. Obtained from relevant sources
   b. Sustained data collection over the relevant time period
   c. Has specified clinical outcomes
   d. Has fully transparent data governance policies, including how the data was collected/extracted, transformed, cleaned, curated, and linked
   e. With data validation, quality assessment, and control measures
3. ***Designs and Analytics***
   The study design should be appropriate to answer the research question, and the data analysis methods should be concrete, rigorous, and compliant with the regulatory standards.[1, 20] Certain analytic methods, such as multivariable models, propensity score analysis, and instrumental variable analysis, can estimate the treatment effects on clinically meaningful outcomes while controlling for the effect of confounders.[16]

2.5.2 How do we evaluate the acceptability of RWE?

There are intrinsic operational and methodological constraints in RWE.[3] These constraints include heterogeneity of data sources, low level of data quality and validity, and potential bias due to unblinded, uncontrolled, or non-randomized treatment allocation.[20]

Despite these limitations, RWE can still be utilized for clinical evaluation of health technologies Different factors should be considered and documented when determining the acceptability of RWE, including the quality and completeness of the data, the

generalizability and acceptability of data, and the timeliness of the evidence.[17, 19, 20] The basic pillars of determining the acceptability of RWE are: [20, 21]

1. ***Transparency and Reproducibility*** – The process by which RWD was generated should be transparent and reproducible. There should be a clear description of the study design and data analysis plan. In studies involving secondary source data, the characteristics of the data source should be clearly described (e.g. type and version of records/registry used, initial purpose of data collection, personnel in charge of collecting and encoding data, setting).
2. ***Validity and accuracy***—The scientific approach used in the RWE should be valid, adhering to standard methodological processes. The study procedure should be sufficiently described. In studies involving secondary source data, methods to ensure data accuracy and completeness, such as data management and quality checks protocols, should be clearly described.

These are general guides on assessing the use and applicability of the RWE in regulatory decision-making. The evaluation may vary depending on the specific clinical question. More detailed discussion on the critical appraisal of RWE is found in Chapter 5.

References

1. National Academies of Sciences, Engineering, and Medicine. Examining the Impact of Real-World Evidence on Medical Product Development. 2018.
2. US FDA. Use of real-world evidence to support regulatory decision-making for medical devices. 2017.
3. Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. Clin Pharmacol Ther 2019;106(1):36.
4. Berger M, Crown W, Daniel G, Eichler HG, Goettsch W, Graff J. Improving Transparency in Non-Interventional Research for Hypothesis Testing-WHY, WHAT, and HOW: Considerations from The Real-World Evidence Transparency Initiative. 2020.
5. Akehurst R, Murphy LA, Sola-Morales O, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Using real-world data in the health technology assessment of pharmaceuticals: strengths, difficulties, and a pragmatic way forward. Value Health 2023;26(4):11-19.
6. de Lusignan S, Crawford L, Munro N. Creating and using real-world evidence to answer questions about clinical effectiveness. J Innov Health Inform 2015;22(3):368-73.
7. Lin L, Ahn J, Bayani D, Chan K, Choiphel D, Isaranuwatchai W, et al. Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in ASIA. [Internet]. 2020 [cited 2024 Feb 17]. Available from: https://hiper.nus.edu.sg/realise-guidance/.
8. Graili P, Guertin JR, Chan KK, Tadrous M. Integration of real-world evidence from different data sources in health technology assessment. J Pharm Pharm Sci 2023;26:11460.
9. National Medical Products Administration. Guiding principles of real world data used to generate real world evidence (trial). 2021.
10. Dang A. Real-world evidence: a primer. Pharm Med 2023;37(1):25-36.
11. Shunmugavelu M, Panda JK, Sehgal A, Makkar BM. A review on generation of real-world evidence. Clin Diabetol 2021;10(5):412-19.
12. Ahn EK. A brief introduction to research based on real-world evidence: Considering the Korean National Health Insurance Service database. Integr Med Res 2022;11(2):100797.
13. Le-Rademacher J, Gunn H, Yao X, Schaid DJ. Clinical trials overview: from explanatory to pragmatic clinical trials. Mayo Clin Proc 2023;98(8):1241-53.
14. Oehrlein EM, Schoch S, Burcu M, McBeth JF, Bright J, Pashos CL, Willke R, Love TR, Mattingly TJ 2nd, Perfetto EM; Patient-Centered Real-World Evidence Working Group. Developing Patient-Centered Real-World Evidence: Emerging Methods Recommendations From a Consensus Process. Value Health. 2023 Jan;26(1):28-38.
15. Ramamoorthy A, Huang SM. What does it take to transform real-world data into real-world evidence? Clin Pharmacol Ther 2019;106(1):10-18.
16. Liu M, Qi Y, Wang W, Sun X. Toward a better understanding about real-world evidence. Eur J Hosp Pharm 2022;29(1):8-11.
17. Bhatt A. Conducting real-world evidence studies in India. Pers Clin Res 2019;10(2):51-56.

18. de Lusignan S, Crawford L, Munro N. Creating and using real-world evidence to answer questions about clinical effectiveness. *J Innov Health Inform*. 2015;22(3):368-373.
19. Miksad RA, Abernethy AP. Harnessing the power of real-world evidence (RWE): a checklist to ensure regulatory-grade data quality. Clin Pharmacol Ther 2018;103(2):202-5.
20. Valla V, Tzelepi K, Charitou P, Lewis A, Polatidis B, Koukoura A, et al. Use of Real-World Evidence for International Regulatory Decision Making in Medical Devices. Int J Digit Health 2023;3(1):1.
21. Fleurence RL, Kent S, Adamson B, Tcheng J, Balicer R, Ross JS, et al. Assessing Real-World Data From Electronic Health Records for Health Technology Assessment: The SUITABILITY Checklist: A Good Practices Report of an ISPOR Task Force. Value Health. 2024 Jun;27(6):692-

Chapter III

When can we use Real-World Evidence for Clinical Evaluation in HTA?

3.1 Introduction

3.2 When are RCTs insufficient to make a decision?

3.3 When are RCTs impossible to conduct due to feasibility or ethical issues?

Chapter Summary

Systematic reviews of RCTs remain the most appropriate evidence for evaluating the clinical efficacy of health technology. RWE may be used in the clinical evaluation of health technologies when RCTs are insufficient to make a decision and RCTs are impossible to conduct due to feasibility or ethical issues

3.1 Introduction

Systematic reviews of RCTs remain the most appropriate evidence for clinical evaluation of health technologies. It is crucial to exercise caution when using RWE for HTA due to inherent biases (e.g., selection bias and confounding bias) and the potential for generating misleading conclusions that may under- or overestimate the effectiveness of health technologies.[1,2] However, RWE may be used in the clinical evaluation of health technologies when RCTs are insufficient to make a decision and RCTs are impossible to conduct due to feasibility or ethical issues.[1,2] These scenarios are discussed further in Sections 3.2 and 3.3.

3.2 When are RCTs insufficient to make a decision?

In performing systematic reviews of RCTs in the clinical evaluation of health technologies, the available RCTs may be deemed 'insufficient' to inform regulatory decision-making when RCTs have insufficient study time frames to capture long-term outcomes. Clinical outcomes such as mortality or health-related quality of life are commonly measured in RCTs in short-term time points. Long-term outcomes are rarely captured in clinical trials due to feasibility issues. In these instances, the use of RWE in regulatory decision-making is focused on continued monitoring of the benefit-risk trade-off after initial approval, with particular attention to long-term safety concerns.[3] Extrapolating the survival curve of patients from short-term trials is also done to estimate long-term clinical outcomes. RWE may be used to supplement but not validate these extrapolated survival curves due to the inherent survival bias in RWE.[4]

Moreover, surrogate outcomes are often measured in RCTs due to faster outcome accrual, leading to trials with short follow-up periods.[5] Common examples include lipid levels as a surrogate outcome for cardiovascular events, blood glucose as a surrogate outcome for diabetic complications, bone density as a surrogate outcome for fracture occurrence, and change in antibody titers as a surrogate outcome for prevention of vaccine-associated illness. However, surrogate outcomes may lead to an overestimate of intervention effect. In certain instances, reliance on surrogate outcomes may lead to an increased risk of harm among patients.[6-7] An example is the use of clofibrate for patients with heart disease. Although clofibrate was shown to reduce cholesterol levels (surrogate outcome), it was later on shown to lead to increased mortality (clinical outcome).[8] RWE may be used to validate surrogate outcomes and establish an association with more relevant clinical outcomes. Several HTA organizations have established methods for using RWE to link surrogate outcomes with clinical outcomes.[5] It is important to note that surrogate outcomes cannot replace clinical outcomes in decision-making.[4]

There are ongoing international initiatives to provide guidance on using surrogate outcomes for decision-making and to strengthen the relationship between surrogate outcomes and long-term health outcomes [9-11]. The outputs of these initiatives will be included in future updates of this methods guide.

3.3 When are RCTs impossible to conduct due to feasibility or ethical issues?

There are instances when it is unethical or not feasible to randomize study participants to a control arm (e.g. oncologic conditions without a standard of care, fatal conditions such as rabies). In these instances, RWE may be used to complement the results of single-arm interventional trials.[12] RWE may also be used to complement results of N-of-1 trials, which are crossover trials wherein participants are given multiple interventions in a randomized order, with each participant serving as his/her own control.[13]

Certain population groups such as those with multiple co-morbidities, geriatric patients, pregnant and lactating women, and children are also commonly not included in RCTs.[4] RWE may provide evidence on the safety and effectiveness of health technologies in these population groups.[14]

Rare diseases are another example where RCTs are almost impossible to conduct due to the small number of patients available for recruitment, variability in clinical presentation and prognosis, difficulty in accurately diagnosing patients, and lack of established standards of care.[4, 15-17] In the Philippines, rare diseases are defined as diseases that occur in 1 in 20,000 Filipinos.[18] These disorders include inherited metabolic disorders and other diseases with similar rare occurrences.[19] There are cases where regulatory agencies accepted RWE to support health technology approval, but mainly in the context of oncology and other rare conditions.[9, 17, 20]

Pandemics are emergency situations warranting urgent decision-making for new health technologies or new indications for existing health technologies. In these exceptional situations, RWE may be used to guide decision-making while results of RCTs are not yet available. However, immediate re-assessment should be done once results of RCTs are available since RCTs are still the most appropriate study design for clinical evaluation of health technologies.

References:

1. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000;342(25):1887-92.
2. Cuello-Garcia CA, Santesso N, Morgan RL, Verbeek J, Thayer K, Ansari MT, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. J Clin Epidemiol 2022;142:200-8.
3. Jansen M, Dekkers O, Cessie S, van der Laan M, Hillege H, van der Graaf R, et al. Real-world evidence to inform regulatory decision making: A scoping review. Clinical Pharmacology & Therapeutics 2024;0.
4. Lin L, Ahn J, Bayani D, Chan K, Choiphel D, Isaranuwatchai W, et al. Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in ASIA. [Internet]. 2020 [cited 2024 Feb 17]. Available from: https://hiper.nus.edu.sg/realise-guidance/.
5. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, et al. Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. Pharmacoeconomics 2020;38(10):1055-70.
6. Ciani O, Buyse M, Garside R, Pavey T, Stein K, Sterne JAC, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. BMJ 2013;346:f457.
7. Rodríguez-Gutiérrez R, Montori VM. Glycemic control for patients with type 2 diabetes mellitus: our evolving faith in the face of evidence. Circ Cardiovasc Qual Outcomes 2016;9(5):504-12.
8. W.H.O. cooperative trial on primary prevention of ischaemic heart disease using clofibrate to lower serum cholesterol: mortality follow-up. Report of the Committee of Principal Investigators. Lancet 1980;2:379 – 85.
9. US FDA. RWD and RWE-focused Demonstration Projects. U.S. Food & Drug Administration, 2023. at https://www.fda.gov/science-research/real-world-evidence/rwd-and-rwe-focused-demonstration-projects.)
10. Baldwin D, Carmichael J, Cook G, Maughan TS, O'Reilly S, Tait D, et al. UK stakeholder perspectives on surrogate endpoints in cancer, and the potential for UK real-world datasets to validate their use in decision-making. Cancer Management and Research 2024;16:791-810.
11. Enhancing Health Technology Assessment Methods: Navigating Surrogate Endpoints and Associated Statistical Standards. ISPOR, 2024. at https://www.ispor.org/heor-resources/presentations-database/presentation/euro2024-3985/18526.)
12. US FDA. Framework for FDA's Real-World Evidence Program. [Internet]. 2018 [cited 2024 Jun 13]. Available from: https://www.fda.gov/drugs/development-resources/advancing-real-world-evidence-program.
13. Hawksworth O, Chatters R, Julious S, Duley L, McKee M, McMahon A, et al. A methodological review of randomised n-of-1 trials. Trials 2024;25.

14. Horton DB, Blum MD, Burcu M. Real-World Evidence for Assessing Treatment Effectiveness and Safety in Pediatric Populations. J Pediatr 2021;238:312-6.
15. Graili P, Guertin J, Chan K, Tadrous M. Integration of real-world evidence from different data sources in health technology assessment. Journal of Pharmacy and Pharmaceutical Sciences. 2023.
16. Dayer V, Drummond M, Dabbous O, Marra C, Kearns B, Hutton J, et al. Real-world evidence for coverage determination of treatments for rare diseases. Orphanet Journal of Rare Diseases 2024;19.
17. Hagen G, Wisløff T. Registry data for use in health technology assessments in Norway – Opportunities and challenges. Norsk Epidemiologi 2021;29:19-27.
18. Department of Health. Integrated Rare Diseases Management Program Strategic Plan 2022-2026. [Internet]. 2022 [cited 2024 Jun 13].
19. Republic of the Philippines. Republic Act No. 10747. An Act Promulgating A Comprehensive Policy In Addressing The Needs Of Persons With Rare Disease. [Internet]. 2015 [cited 2024 Jun 13].
20. Lau C, Jamali F, Loebenberg R. Health Canada usage of real world evidence (RWE) in regulatory decision making compared with FDA/EMA usage based on publicly available information. J Pharm Pharm Sci 2022;25:227-36.

Chapter IV

How do we conduct a systematic search and selection for RWE?

4.1 Introduction

4.2 What are the steps in conducting a systematic search for RWE?

- 4.2.1 Define the research question
- 4.2.2 Perform a systematic literature search
- 4.2.3 Select studies to be included

Chapter Summary

Performing a systematic review of RWE follows the general steps in conducting systematic reviews. The steps in conducting a systematic search for RWE include clearly defining the research question using the PICO (Population, Intervention/Exposure, Comparator, Outcome) framework, performing a systematic literature search, and selecting studies to be included.

4.1 Introduction

Performing a systematic review of RWE follows the general steps in conducting systematic reviews. These general steps are found in the Cochrane Handbook of Systematic Reviews of Interventions[1] and the Philippine HTA Methods Guide (Refer to section 2.2.4.1.1.Evidence Synthesis from Clinical Studies on the Efficacy, Effectiveness and Safety of the Health Technology).[2] A protocol must be developed prior to the conduct of the systematic review to ensure that the process is structured, transparent and reproducible.

This chapter introduces the essential considerations in conducting a systematic literature search specific for RWE, and the selection of RWE to be included in the systematic review. Systematic literature search is a structured process of identifying all relevant studies on a particular topic or research question. This process is important to minimize bias, ensure reproducibility, and provide a solid foundation for the systematic review.[3] When searching for RWE, the challenge often lies in the vast and heterogeneous nature of the data sources.[4] Therefore, a well-structured search strategy is critical to capture the breadth of available evidence while maintaining focus and relevance.

4.2 What are the steps in conducting a systematic search for RWE?

4.2.1 Define the research question

As in any systematic review, the first step is to clearly define the research question using the PICO (Population, Intervention/Exposure, Comparator, Outcome) framework (Table 3). This helps identify the key concepts and terms to be used in the systematic literature search.[5]

Table 3. Elements of the research question

Component	Description	Example
Population (P)	● Refers to the specific group of individuals with a health condition who are likely to benefit from the introduction of the new technology. ● In RWE, this would often include patients with diverse clinical or demographic characteristics, or patients with specific health conditions such as those with rare diseases. ● RWE often fills the gaps in knowledge on populations that are usually excluded in clinical trials such as high risk patients, pediatric or geriatric patients, patients with multiple comorbidities, and ethnic minorities.[6]	In a study examining the effectiveness of a new diabetes management program, the population may be defined as "*adults aged 18-80 with type 2 diabetes in primary care settings*."
Intervention (I) or Exposure (E)	● Refers to the treatment, procedure, or exposure being investigated. ● In RWE, this would often reflect real-world clinical practices which may include treatments with different dosing regimens, surgical procedures, use of medical devices, sequential therapies, and lifestyle interventions.[6]	The intervention in the diabetes management program study may be "*a comprehensive lifestyle modification program that includes diet, exercise, and regular follow-up consultations*."
Comparison (C)	● Refers to the alternative of the intervention, which could be a different treatment, placebo, or standard care. ● In RWE, the most appropriate comparison is standard care. ● A formal comparison group may not be present in certain types of RWE. Pre-post comparison within the same population (also called self-controlled methods) may be done.[7] Single arm trials may also use external comparators (also called historical comparators, external control, or synthetic control).[8]	The comparison group in the diabetes management program study may be "*patients receiving standard diabetes care without the additional lifestyle modification program*."
Outcome (O)	● Refers to the effects of the intervention/exposure that the study aims to measure. ● In RWE for clinical evaluation of health technologies, these can include health utilization outcomes (hospitalization rates, emergency department visits), patient reported outcomes (quality of life, functional status, patient satisfaction), safety outcomes (adverse events, long term safety, withdrawal rates),[9] and even rare outcomes (development of rare complications).[6]	Outcomes in the diabetes management program study may include "*patient-reported quality of life, hospital admissions due to diabetes complications, and self-reported hypoglycemic symptoms.*"

4.2.2 Perform a systematic literature search

When performing literature search, it is essential to search all relevant electronic medical databases to retrieve all relevant studies. Searching only one database (e.g. MEDLINE) may result in missed relevant articles that may have been retrieved through a broader, more comprehensive search across multiple databases.[1] In addition to electronic databases, RWE may also often be found in other information sources. The list of information sources are shown in Table 4.

Table 4. Information sources when conducting systematic literature search for RWE[11]

Source	Description
Bibliographic databases	● Searching more than two databases is important to ensure completeness of the search. ● The most common databases are MEDLINE, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science
Clinical Trial Registries	● Ongoing and completed pragmatic trials and observational studies may be identified in clinical trial registries. ● The most common registries are ClinicalTrials.gov, European Union Clinical Trials Registry, and WHO International Clinical Trials Registry Platform
Health Technology Assessment (HTA) Reports	● Relevant RWE may be identified from the HTA reports of government agencies such as the National Institute for Health and Care Excellence (NICE), Canada's Drug Agency (CDA), and National Health Service Economic Evaluation Database (NHS EED)
Patient-Reported Outcome Measures Databases	● Examples include Patient-Reported Outcomes Measurement Information System (PROMIS) and National Health Service Digital Patient-Reported Outcomes Measurement (NHS PROMs)
Grey literature	● Technical or research reports, doctoral dissertations, and conference papers not indexed in traditional databases may be found in OpenGrey.eu, Health-Related Grey Literature guide, Health Management Information Consortium (HMIC) Database, United States National Technical Information Service (US NTIS), American Psychological Association (APA) PsycExtra, and National Institute for Health Research (NIHR) Journals Library. ● Conference proceedings and abstracts may also be found in the websites and online databases of relevant medical societies

An efficient literature search strategy for each information source is needed to ensure the retrieval of all relevant studies. It is not necessary for the search to be independently conducted by two people; however, the search strategy must be

peer-reviewed prior to conducting the search.[1, 13] The steps in conducting an efficient literature search are summarized in Table 5.

Table 5. Steps in conducting an efficient literature search for RWE [1,2]

Step	Description
1. Identify the PICO key concepts	● Identify the key concepts of the PICO of your research question ● Generate synonyms, related terms, and variations for the P, I, C and O.
2. Identify the M (methodologic filter)	● A methodologic filter is a search strategy designed to retrieve a particular study design. [11] ● You will need to identify which type of study design will be most appropriate for your research question. Since the types of study designs of RWE are highly variable (see chapter II), you may need to use multiple methodologic filters. ● If multiple study designs are appropriate for your research question, you will need to consider the hierarchy of evidence among the observational study designs and prioritize accordingly (i.e. cohort study > case control study > cross-sectional study). ● Validated search filters of the *Scottish Intercollegiate Guidelines Network* are available to improve the efficiency of relevant article retrieval.[5] Alternatively, you may opt to use exclusion criteria as methodologic filter to exclude RCTs, case reports and case series, clinical practice guidelines, narrative reviews, and other study designs that are not applicable to the research question [6] (See Annex 1). ● If there are only a few articles that were retrieved, you may opt not to use any methodologic filter at all, since the study designs of RWE are broad.
3. Prioritize the elements (PICOM) from most to least important	● Prioritization of the PICOM helps create a search strategy that maximizes relevant study retrieval while minimizing the retrieval of irrelevant studies. ● Determine the order in which you enter search terms. The most critical concepts (which should be entered first) are typically the population (P) and intervention/exposure (I/E). ● Outcomes (O) are generally the least important and should not be used as search terms, as they can significantly limit search results.
4. Expand each element sequentially	● Start with the most important concept identified in step 3 and enter all relevant synonyms in the search engine. Combine each synonym using the Boolean operator “OR”. ● Utilize Medical Subject Headings (MeSH) in PubMed, Emtree in Embase, and other controlled vocabularies to find standardized concept terms. Similarly, combine the MeSH terms with the previously identified terms using the Boolean operator “OR” ● You may employ truncation and wildcards to shorten this process. ○ Truncation involves using a symbol (commonly an asterisk *) to replace word endings, capturing all possible variations of a root word. Example-- "epidemiol*" will retrieve "epidemiology," "epidemiological," "epidemiologist," etc.

	○ Wildcards replace a single character within a word to account for different spellings. Example-- "Wom*n" retrieves both "woman" and "women."
5.Intersect the expanded element sequentially	● Sequentially combine the synonyms for each element using the Boolean operator "AND" to narrow down the search to relevant studies.
6. Examine the yield for miss-hits and misses	● Miss-hits are irrelevant articles retrieved in the search. This is commonly caused due to broad search terms that capture unrelated studies or the use of terms with multiple meanings. ● Misses are relevant studies that were not retrieved in the search. This is commonly caused by using too narrow or specific search terms, overuse of filters or limits, or omitting synonyms. ● If your initial search results yield many miss-hits and misses, modify your search strategy accordingly. ○ To reduce miss-hits, add specific terms or qualifiers, or use NOT operators (e.g. diabetes NOT type 1). ○ To capture misses, expand search terms to include more synonyms and review reference lists of key articles for additional studies.
7. Revise the search strategy and re-run the search	● Modify the search strategy based on the findings from your review and run the search again.

The following should be assessed during peer review of the search strategy: a) correct translation of the research question to the PICO elements, b) appropriate selection of the methodologic filter, c) appropriate use of Boolean operators, and d) inclusion of relevant free text and subject headings.[14] Consultation with subject matter experts may also be done to validate the search strategy and ensure that all relevant studies are captured.[7]

4.2.3 Select studies to be included

Identifying the PICOM in your research question also facilitates the screening and selection of relevant RWE. Identifying the PICOM elements of your research question can quickly show if the RWE study that was retrieved during your search is suitable for inclusion into your review.

The selection process starts with the screening of titles and abstracts to exclude studies that are clearly irrelevant. Next, retrieve and review the full texts of potentially relevant studies. The pre-specified inclusion and exclusion criteria in the systematic review are applied to determine final eligibility of articles. At least two independent reviewers should screen each study to reduce bias and increase reliability. Any disagreements should be resolved through discussion or by consulting a third reviewer. It is also important to note whether multiple reports come from the same study.[11]

Document the selection process using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram to document the number of studies identified, screened, excluded, and included in the review.[15] Provide clear reasons for

studies excluded at the full-text screening stage. Lastly, web-based systematic review support tools such as Covidence, DistilleSR, and Abstrackr can be used to facilitate the screening process.

References:

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ WV (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane Handbook for Systematic Reviews of Interventions. doi:10.1002/9780470712184.fmatter
2. Health Technology Assessment Unit, Department of Health Philippine HTA Methods Guide. 1st ed. 2020
3. Briere JB, Bowrin K, Taieb V, Millier A, Toumi M, Coleman C. Meta-analyses using real-world data to generate clinical and epidemiological evidence: a systematic literature review of existing recommendations. *Curr Med Res Opin*. 2018;34(12):2125-2130. doi:10.1080/03007995.2018.1524751
4. Shunmugavelu MS, Panda JK, Sehgal A, Makkar BM. A review on generation of real-world evidence. *Clin Diabetol*. 2021;10(5):412-419. doi:10.5603/DK.2021.0049
5. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:1-6. doi:10.1186/1472-6947-7-16
6. Graili P, Guertin JR, Chan KKW, Tadrous M. Integration of real-world evidence from different data sources in health technology assessment. J Pharm Pharm Sci. 2023;26(July):11460. doi:10.3389/jpps.2023.11460
7. Gokhale M, Stürmer T, Buse JB. Real-world evidence: the devil is in the detail. *Diabetologia*. 2020;63(9):1694-1705. doi:10.1007/s00125-020-05217-1
8. Patel D, Grimson F, Mihaylova E, Shajahan Y, Johansen M, Tzeng J, et al. Use of External Comparators for Health Technology Assessment Submissions Based on Single-Arm Trials. *Value Heal*. 2021;24(8):1118-1125. doi:10.1016/j.jval.2021.01.015
9. Kc S, Lin LW, Bayani DBS, Ahn J, Tan Y, Lee Y, et al. What, Where, and How to Collect Real-World Data and Generate Real-World Evidence to Support Drug Reimbursement Decision-Making in Asia: A reflection into the Past and A Way Forward. *Int J Heal Policy Manag*. 2023;12(1):6858. doi:10.34172/ijhpm.2023.6858
10. Wang S V., Schneeweiss S. Assessing and Interpreting Real-World Evidence Studies: Introductory Points for New Reviewers. *Clin Pharmacol Ther*. 2022;111(1):145-149. doi:10.1002/cpt.2398
11. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane handbook for systematic reviews of interventions. Wiley; 2019. doi:10.1002/9781119536604.
12. Ayiku L, Hudson T, Williams C, Levay P, Jacob C. The NICE OECD countries' geographic search filters: Part 2—validation of the MEDLINE and Embase (Ovid) filters. *J Med Libr Assoc*. 2021;109(4):583-589. doi:10.5195/jmla.2021.1224
13. Franco JVA, Garrote VL, Escobar Liquitay CM, Vietto V. Identification of problems in search strategies in Cochrane Reviews. *Res Synth Methods*. 2018;9(3):408-416. doi:10.1002/jrsm.1302

14. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016 Jul;75:40-6.
15. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021; 372:n71 doi:10.1136/bmj.n7

Chapter V

What are the appropriate methods for critical appraisal of available RWE?

5.1 Why do we need to perform a critical appraisal of RWE?

5.2 What are the sources of bias in RWE?

5.3 Which appraisal tool do we use for assessing risk of bias?

5.4 What are the common issues in the reporting of RWE?

5.5 Which reporting guideline do we use when assessing RWE?

Chapter Summary

The usefulness of RWE is often questioned due to concerns on quality. Appraising the quality of RWE is critical for regulatory decision-making. The sources of bias in RWE depend on the study design used and may include confounding, selection, performance, attrition, detection and reporting bias. The recommended appraisal tools are Cochrane Risk of Bias Tool version 2 for pragmatic clinical trials, and Risk Of Bias In Non-randomised Studies - of Interventions or Exposures (ROBINS-I or ROBINS-E) for observational studies with primary or secondary source data. Common issues in the reporting of RWE are Insufficient and selective reporting. The recommended reporting guidelines for assessing RWE are the modified Consolidated Standards of Reporting Trials (CONSORT) checklist for pragmatic clinical trials, Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for observational studies with primary source data, and Reporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist for observational studies with secondary source data.

5.1 Why do we need to perform a critical appraisal of RWE?

Despite the increasing interest and uptake of RWE by regulatory agencies and other policymakers[1, 2], the usefulness of RWE is often questioned due to concerns regarding their quality. These concerns involve (a) risk of bias of RWE and (b) unclear or inadequate reporting of key study parameters.[3] Thus, appraising the quality of RWE included in the clinical evaluation of health technologies is critical for regulatory decision-making. Similar to the selection of studies, at least two independent reviewers should perform a critical appraisal of the included RWE to increase reliability. Any disagreements should be resolved through discussion or by consulting a third reviewer.

There is no definitive, standard recommendation on how to best appraise the risk of bias and reporting quality in RWE. This limitation may be attributed to the variability in the definition, data sources, and study designs in RWE.[4] Furthermore, there are various tools that can be used to assess the risk of bias and reporting quality in RWE, with no consensus on how to select the best appraisal tool.[5-7] Although societies, regulators, and decision-makers recommend the use of standard reporting guidelines for RWE, these largely remain unused.[6] There is low adoption of reporting guidelines due to the diverse study designs in RWE necessitating the use of varying reporting guidelines.[8, 9]

5.2 What are the sources of bias in RWE?

The sources of bias in RWE depend on the study design used. Pragmatic clinical trials (PCTs) evaluate the effectiveness of interventions outside highly controlled settings by incorporating real-world elements into clinical trial design. Although randomization is maintained in PCTs, real-world elements including heterogeneous populations, less-standardized treatment protocols, and delivery of treatments in routine clinical settings are incorporated into the study design. There are inherent challenges in PCTs due to possible bias from confounding factors, non-adherence to treatment, failure to achieve an absolutely equal distribution of differences by randomization, and lack of double-blinding.[11,12]

Observational studies are prone to bias due to confounding, selection of participants, detection of effects, or attrition.[10] The precision of the effect estimates may also be affected by an inadequate sample size (i.e., insufficient power to test study hypothesis) or lack of study efficiency (i.e., absence of stratification).[13] The different sources of bias in RWEs are summarized in Table 6 below.

Table 6. Sources of bias in real-world evidence

Bias	**Definition**	**Examples**
Confounding bias [14-23]	Systematic distortion in the measure of association/effect between the exposure/intervention and outcome that arise when patient prognostic characteristics, such as disease severity or comorbidity, influence both the exposure/intervention and outcomes Note: It is ideal that reviewers pre-specify the confounding variables relevant to their research question.	• Baseline confounding (i.e., when one or more baseline prognostic variables predict whether an individual receives the intervention/exposure or the comparator) • Time-varying confounding (i.e., when individuals switch between intervention/exposure and the comparator, or when post-baseline prognostic factors affect the intervention/exposure received) • No statistical adjustments made to account for confounding (i.e. no inverse probability weighting or regression analysis done) • Confounding by indication (i.e., physician-directed selection or self-selection of treatment)
Selection bias [14-17, 19-23]	Bias that arises when the study participants are not representative of the target population to which the conclusions will apply.	• Non-random allocation of study participants (for pragmatic trials) • No/inadequate allocation concealment (for pragmatic trials) • Incidence-prevalence bias (e.g., exclusion of individuals with severe or mild disease) • Inappropriate selection of controls • Non-response bias (i.e., when non-responders from a sample differ in a meaningful way to responders; common in surveys) • Healthy entrant effect (i.e. Participants who consent to be part of a research study may be have better health conditions or health-seeking behavior compared to the general population)
Performance bias [14-17, 19-23]	Systematic differences in the care provided to participants in the intervention/exposure and comparator groups other than the intervention under investigation (for studies with at least 2 groups) or	• Variation in delivery or non-adherence to the intended intervention • Contamination bias (i.e., when study participants in one group are exposed to the intervention that is intended for the other group)

	within the group (for single-arm studies)	• Difference in co-interventions
Attrition bias [14, 16-17, 20-23]	Bias that arises due to the loss of participants from the study and how they were accounted for in the results	• Incomplete follow-up • Differential loss to follow-up due to prognostic factors • Exclusion of individuals with missing data
Detection bias [14-17, 19-23]	Systematic differences in outcome assessment among the groups being compared (for studies with at least 2 groups) or within the group being studied (for single-arm and cross-sectional studies)	• Misclassification of the exposure or intervention, covariates, or outcomes because of variable definitions and timings • Absent/inadequate blinding of outcome assessor • Using different methods to assess outcomes among patient groups • Faulty measurement techniques • Recall bias (i.e., inaccurate recall of information from memory) • Lead-time bias (i.e., when screening interventions result in earlier diagnosis) • Immortal time bias (i.e., when the follow-up includes a period where participants in the exposed group cannot experience the outcome)
Reporting bias [20-23]	Systematic differences between reported and unreported findings, or selective reporting of results depending on the available findings	• Differential or incomplete reporting of outcomes • Analysis that is not in accordance with prespecified plan / study protocol

5.3. Which appraisal tool do we use for assessing risk of bias?

Existing appraisal tools for RWE vary in terms of focus, level of detail, signalling questions and response options. Regardless of which tool is used, it is important for the reviewer to consider the following when appraising RWE: [7, 24]

a. Use the most appropriate tool to assess risk of bias depending on the study design of the RWE
b. Assess the generalizability to the local population to which the RWE will be applied for regulatory decision-making
c. Check the study registration, availability of study protocol, and adherence of RWE to reporting guidelines

Upon systematic review of literature, there are 49 eligible appraisal tools for non-randomized studies of interventions involving real-world data. However, only a few of these appraisal tools were found to comprehensively assess the items for methodological

quality. [5] These include the Research Triangle Institute Item Bank (RTI Item Bank) [13] and Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) [25], as well as tools for specific research areas such as PROTECT checklist for drug safety studies [26] and MINORS checklist for surgical research.[27] A dedicated appraisal tool for RWE using EHR, called Assessment of Real-World Observational Studies (ArRoWs) is also currently undergoing validation.[26]

After careful review of all appraisal tools and literature related to its use, the **Cochrane Risk of Bias Tool version 2 [RoB 2]** is recommended for assessing risk of bias in pragmatic clinical trials.[10] For observational studies that generate primary source data, the **ROBINS-I** tool for studies of interventions, or the **ROBINS-E** tool for studies of exposures is recommended. Secondary source data are typically analyzed using observational study designs. The ROBINS-I and ROBINS-E tool may be used for observational studies using secondary source data as well. (Table 7).[25, 29] In recent years, target trial emulation has also emerged as a novel approach wherein principles of randomized trials are applied to existing data sources to emulate an open-label RCT.[30-32] There are currently no standard appraisal tools for target trial emulation, but some authors use an adaptation of the ROBINS-I tool.[33]

Table 7. Recommended tools for appraising risk of bias in RWE

Study Design	**Appraisal Tool**	**Comments**
Pragmatic clinical trials	Cochrane Risk of Bias Tool version 2 [RoB 2]	● Although no specific risk of bias tool has been developed specifically for pragmatic clinical trials, RoB 2 has been used in published systematic reviews for appraising pragmatic clinical trials.[34-36] ● Using ROB2 may offer greater consistency in situations where both RCTs and PCTs are included in the same evidence synthesis.
Observational studies	Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) or Risk Of Bias In Non-randomised Studies - of Exposures (ROBINS-E)	● Although completion of the ROBINS-I or ROBINS-E tool may take considerable time, it shares many features with other risk of bias tools (e.g.,. RoB 2, QUADAS-2) and is the recommended tool by the Cochrane Scientific Committee for systematic reviews

5.4 What are the common issues in the reporting of RWE?

Insufficient and selective reporting of how RWEs are generated prevents decision-makers from making comprehensive and balanced evidence summaries. [37]

Moreover, significant variability in the format, completeness, and methods for RWE have been noted across countries.[1, 2] Using standardized reporting guidelines would enhance transparency of RWE, facilitate planning and generation of high-quality RWE, and increase the confidence of stakeholders in using these data sources for HTA.

Common deficiencies in reporting of RWE have been identified in a 2023 systematic review of 54 RWE studies.[6] We urge reviewers to pay attention to the following items that are usually either unreported or only partially reported:

a. Use of a reporting guideline (e.g., STROBE, RECORD-PE) [38]
b. Protocol registration details
c. Target study population, time horizon, and setting
d. Information on confounding variables
e. Data analysis specifications (e.g.,. confounding adjustment method, propensity score matching or weighting, missing data, subgroup analysis)
f. Sensitivity analyses (e.g., rationale, strengths and weakness compared to primary analysis)
g. Criteria for defining index date (i.e., date or calendar time range when subjects enter the study population cohort)
h. Follow-up and outcome data (e.g., washout window, dates of outcome measurement)
i. Metadata about data source and software (e.g., data linkage, data conversion, software to create study population)

5.5. Which reporting guideline do we use when assessing RWE?

Many frameworks and tools have been developed to standardize reporting of key study components of RWE. The **modified Consolidated Standards of Reporting Trials (CONSORT) checklist** (Annex 2) is recommended for assessing pragmatic clinical trials. This checklist was adapted primarily from the extension of the CONSORT statement for pragmatic trials [39] and the February 2024 reporting template created by the National Institutes of Health Pragmatic Trials Collaboratory.[40]

For observational studies generating primary source data, the **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist** is recommended.[41] If data for the observational study was obtained from routinely collected health data or secondary source data (e.g., disease registries, health administrative data, electronic health records, epidemiological surveillance data, claims database), the **Reporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist** [42] or its extension for pharmacoepidemiologic research (**RECORD-PE**) is recommended.[43] These reporting guidelines are detailed in Annex 3. For all reporting checklists, each item may be rated as either *reported, partially reported, or not reported*. No specific reporting checklists are established for target trial emulation.[33] Commonly used checklists include STROBE, ISPOR Good Research Practices for Comparative Effectiveness Research, RECORD and RECORD-PE.[44]

Table 8. Recommended guidelines for assessing reporting quality of RWE

Study Design	**Reporting guideline**	**Comments**
Pragmatic clinical trials	Modified Consolidated Standards of Reporting Trials (CONSORT) checklist	● Based primarily on the extension of the CONSORT statement for pragmatic trials[35] ● Contains additional items from the February 2024 reporting template created by the NIH Pragmatic Trials Collaboratory to account PCTs where the data sources is from clinical or administrative databases instead of a dedicated research database
Observational studies that generate primary source data	Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist	● STROBE checklist for observational studies where data was collected in accordance to specific pre-defined research objectives[41]
Observational studies using secondary source data	REporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist or REporting of studies Conducted using Observational Routinely-collected Data for PharmacoEpidemiology (RECORD-PE) checklist	● RECORD checklist for observational studies using routinely collected health data without specific pre-defined research objectives[42] ● RECORD-PE checklist for pharmacoepidemiologic research using routinely collected health data[43]

References

1. Kc S, Lin LW, Bayani DBS, Ahn J, Tan Y, Lee Y, et al. What, Where, and How to Collect Real-World Data and Generate Real-World Evidence to Support Drug Reimbursement Decision-Making in Asia: A reflection into the Past and A Way Forward. *Int J Heal Policy Manag*. 2023;12(1):6858. doi:10.34172/ijhpm.2023.6858
2. Thokagevistk K, Coppo C, Rey L, Luyten J, Boulanger D, Tzeng J, et al. Real-world evidence to reinforce clinical trial evidence in health technology assessment: A critical review of real-world evidence requirements from seven countries and recommendations to improve acceptance. Journal of Market Access & Health Policy. 2024;12(2), 105-117. doi:10.3390/jmahp12020009
3. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, Higgins JPT. Chapter 25: Assessing risk of bias in a non-randomized study. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.*), Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023)*. Cochrane. 2023. Retrieved from www.training.cochrane.org/handbook
4. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W, van der Wilt GJ, et al. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value in Health. 2017; *20*, 858–865.
5. Jiu L, Hartog M, Wang J, Kooijman M, Hillege H, Goettsch W, et al. Tools for assessing quality of studies investigating health interventions using real-world data: A literature review and content analysis. BMJ Open. 2024;14(2), e075173. doi:10.1136/bmjopen-2023-075173
6. Khambholja K, Gehani M. Use of structured template and reporting tool for real-world evidence for critical appraisal of the quality of reporting of real-world evidence studies: A systematic review. Value in Health. 2023;26(3), 427-434. doi:10.1016/j.jval.2022.09.003
7. Briere JB, Bowrin K, Taieb V, Millier A, Toumi M, Coleman C. Meta-analyses using real-world data to generate clinical and epidemiological evidence: A systematic literature review of existing recommendations. Current Medical Research and Opinion. 2018;34(12), 2125-2130. doi:10.1080/03007995.2018.1524751
8. Khambholja K, Gehani M. Use of Structured Template and Reporting Tool for Real-World Evidence for Critical Appraisal of the Quality of Reporting of Real-World Evidence Studies: A Systematic Review. Value Health. 2023 Mar;26(3):427-434. doi: 10.1016/j.jval.2022.09.003. Epub 2022 Oct 7. PMID: 36210293.
9. Wang H, Chen Y, Lin Y, Abesig J, Wu IX, Tam W. The methodological quality of individual participant data meta-analysis on intervention effects: systematic review. BMJ. 2021 Apr 19;373:n736. doi: 10.1136/bmj.n736. PMID: 33875446; PMCID: PMC8054226.
10. Godwin M, Ruhland L, Casson, I., van der Windt D, Weller D, Little P, et al. Pragmatic controlled clinical trials in primary care: The struggle between external and internal validity. BMC Medical Research Methodology. 2003;*3*, 28. doi:10.1186/1471-2288-3-28
11. Ford I, Norrie J. Pragmatic Trials. The New England Journal of Medicine. 2016;375(5), 454-463. doi:10.1056/NEJMra1510059

12. Sterne JAC, Savović J, Page MJ, Elbers RG, Reed PT, Hutton B, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366, l4898. doi:10.1136/bmj.l4898
13. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. Journal of Clinical Epidemiology. 2012;65(2), 163-178. doi:10.1016/j.jclinepi.2011.05.008
14. Gliklich R, Leavy M. Assessing real-world data quality: The application of patient registry quality criteria to real-world data and real-world evidence. Therapeutic Innovation & Regulatory Science. 2020;54(2):303-307.
15. Gokhale M, Stürmer T, Buse J. Real-world evidence: the devil is in the detail. Diabetologia. 2020;63(9):1694-1705.
16. Center for Medical Technology Policy. A Practical Tool for Assessing Relevance and Rigor of Real World Evidence: User's Guide. RWE Decoder Too. Baltimore, Maryland: CMTP. 2017.
17. Taur S. Observational designs for real‐world evidence studies. Perspect Clin Res. 2022;13:12-16.
18. Rosen E, Ritchey M, Girman C. Can weight of evidence, quantitative bias, and bounding methods evaluate robustness of real-world evidence for regulator and health technology assessment decisions on medical interventions? Clinical Therapeutics. 2023;45:1266-1276.
19 Lin L, Ahn J, Bayani D, Chan K, Choiphel D, Isaranuwatchai W, et al. Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in ASIA. 2020. https://hiper.nus.edu.sg/realise-guidance/. Accessed February 17, 2024.
20. Canada's Drug and Health Technology Agency, Health Canada. Guidance for Reporting Real-World Evidence. CADTH Methods and Guidelines. 2023. https://www.cadth.ca/guidance-reporting-real-world-evidence
21. Institut national d'excellence en santé et en services sociaux (INESSS). Integration of Real-World Data and Evidence to Support Decision-Making in the Pharmaceutical Sector. 2022. https://www.inesss.qc.ca/en/publications/publications/publication/integration-of-real-world-data-and-evidence-into-assessments-to-support-decision-making-in-the-pharmaceutical-sector.html.
22. Haute Autorité de santé. Real-world studies for the assessment of medicinal products and medical devices. 2021. https://www.has-sante.fr/jcms/p_3284524/en/real-world-studies-for-the-assessment-of-medicinal-products-and-medical-devices.
23. National Institute for Health and Care Excellence. NICE real-world evidence framework. NICE. www.nice.org.uk/corporate/ecd9. Published 2022. Accessed February 24, 2024.
24. Curtis LH, Sola-Morales O, Heidt J, Kearns B, Tan Y, Hutton J, et al. Regulatory and HTA considerations for development of real-world data derived external controls. Clinical Pharmacology & Therapeutics. 2023;114(2), 303-315. doi:10.1002/cpt.2913
25. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Campbell B, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355, i4919. doi:10.1136/bmj.i4919

26. Faillie JL, Ferrer P, Gouverneur A, Durrieu G, Bouzillé G, Ménard J, et al. A new risk of bias checklist applicable to randomized trials, observational studies, and systematic reviews was developed and validated to be used for systematic reviews focusing on drug adverse events. Journal of Clinical Epidemiology. 2017;*86*, 168-175. doi:10.1016/j.jclinepi.2017.04.023
27. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): Development and validation of a new instrument. The Australian and New Zealand Journal of Surgery. 2003;73(9), 712-716. doi:10.1046/j.1445-2197.2003.02748.x
28. Coles B, Tyrer F, Hussein H, Dhalwani N, Khunti K. Development, content validation, and reliability of the Assessment of Real-World Observational Studies (ArRoWS) critical appraisal tool. Annals of Epidemiology. 2021;*55*:57-63.e15. doi:10.1016/j.annepidem.2020.09.014
29. Higgins JPT, Morgan RL, Rooney AA, Taylor KW, Thayer KA, Silva RA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environment International. 2024; 186:108602. doi:10.1016/j.envint.2024.108602
30. Honap S, Danese S, Peyrin-Biroulet L. Target Trial Emulation: Improving the Quality of Observational Studies in Inflammatory Bowel Disease Using the Principles of Randomized Trials. Inflammatory Bowel Diseases. 2024.
31. Bray B, Ramagopalan S. R WE ready for reimbursement? A round up of developments in real-world evidence relating to health technology assessment: part 14. Journal of Comparative Effectiveness Research 2023:e230189.
32. Moler-Zapata S, Hutchings A, O'Neil S, Silverwood R, Grieve R. Emulating target trials with real-world data to inform health technology assessment: Findings and lessons from an application to emergency surgery. Value Health 2023;26:1164-74.
33. Bigirumurame T, Hiu SKW, Teare MD, Wason JMS, Bryant A, Breckons M. Current practices in studies applying the target trial emulation framework: a protocol for a systematic review. BMJ Open. 2023;13:e070963.
34. Breet E, Matooane M, Tomlinson M, Bantjes J. Systematic review and narrative synthesis of suicide prevention in high-schools and universities: A research agenda for evidence-based practice. BMC Public Health. 2021; 21(1):1116. doi:10.1186/s12889-021-11124-w
35. Maurer LF, Schneider J, Miller CB, Espie CA, Kyle SD. The clinical effects of sleep restriction therapy for insomnia: A meta-analysis of randomised controlled trials. Sleep Medicine Reviews. 2021;58:101493. doi:10.1016/j.smrv.2021.101493
36. Fitzpatrick C, Gillies C, Seidu S, Misso K, Khunti K, Majeed A, et al. Effect of pragmatic versus explanatory interventions on medication adherence in people with cardiometabolic conditions: A systematic review and meta-analysis. BMJ Open. 2020; 10(7):e036575. doi:10.1136/bmjopen-2019-036575
37. Orsini LS, Berger M, Crown W, D'Amico R, Haller D, McGowan J, et al. Improving transparency to build trust in real-world secondary data studies for hypothesis testing-Why, what, and how: Recommendations and a road map from the Real-World Evidence Transparency Initiative. Value in Health. 2020;23(9):1128-1136. doi:10.1016/j.jval.2020.04.002

38. Sterrantino A. Observational studies: practical tips for avoiding common statistical pitfalls. The Lancet Regional Health Southeast Asia. 2024;25.
39. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Bossuyt PM, Clarke M, et al. Improving the reporting of pragmatic trials: An extension of the CONSORT statement. BMJ. 2008;337, a2390. doi:10.1136/bmj.a2390
40. NIH Pragmatic Trials Collaboratory Coordinating Center. Reporting pragmatic clinical trials. 2024. Retrieved June 10, 2024, from https://dcricollab.dcri.duke.edu/sites/NIHKR/KR/PCT%20Reporting%20Template.pdf.
41. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. International Journal of Surgery. 2014; 12(12):1500-1524. doi:10.1016/j.ijsu.2014.07.014.
42. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Peter J, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Medicine. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885.
43. Langan SM, Schmidt SA, Wing K, Mallaiah S, Hviid A, Bøtker HE, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363:k3532. doi:10.1136/bmj.k3532.
44. Hansford HJ, Cashin AG, Jones MD, Wiggins J, McKenzie S, McDonald S, et al. Reporting of observational studies explicitly aiming to emulate randomized trials: a systematic review. JAMA Netw Open. 2023;6:e2336023.

Chapter VI

How do we extract data from RWE?

6.1 What are the appropriate methods for data extraction of RWE?

6.2 How do we extract general study information?

6.3 How do we extract the elements of the research question?

6.4 How do we extract key components of the study methodology?

- 6.4.1 Study design
- 6.4.2 Inclusion and Exclusion criteria
- 6.4.3 Data Source
- 6.4.4 Data Collection Methods
- 6.4.5 Data Analysis Methods

6.5 How do we extract the results reported in RWE?

Chapter Summary

The process of data extraction should be in accordance with the pre-defined protocol. The general study information to be extracted include title, author(s), year of publication, journal name, volume and issue number, country or geographical region, coverage period, and specific data source. The elements of the research question of each study should be identified, following the PICO framework. Key components of the study methodology to be extracted include study design, inclusion and exclusion criteria, data source, data collection methods, data analysis methods. For dichotomous and time-to-event outcomes, it is recommended to extract the adjusted effect estimates (adjusted odds ratio, adjusted relative risk or adjusted hazard ratio) and its standard error or 95% confidence interval. For continuous outcomes, it is recommended to extract the partial regression coefficients derived from the regression model.

6.1 What are the appropriate methods for data extraction of RWE?

Data extraction entails collecting pertinent information from studies that meet the eligibility criteria to facilitate data synthesis and allow clinical evaluation of health technologies. The data extraction process should be in accordance with the pre-defined protocol to ensure a systematic and reproducible process.[1] At least two independent reviewers should perform data extraction to increase reliability. Any disagreements should be resolved through discussion or by consulting a third reviewer.

The following information may be extracted from each RWE: general study details (title, author, year, journal name, setting), elements of the research question (PICO), study methodology components (study design, eligibility criteria, data source, data collection, and data analysis methods), and study results. A template for a data extraction form is provided at Annex 4 and is designed to assist you in creating your customized version. You will need to adapt this template to suit your specific research question.

6.2 How do we extract general study information?

The general study information typically collected for RWE includes several key items, such as title of the study, author(s), year of publication, journal name, and the volume and issue number are recorded. The DOI or URL is also obtained to ensure easy access to the study. For studies with primary source data, it is important to note the country or geographical region where the study was conducted, and the coverage period including the time period and length of observations in data collection. For studies involving secondary source data, it is important to note the country or geographical region where the data was collected, the time period of data collection, and the specific data source (e.g. registry, electronic medical records, databases).

6.3 How do we extract the elements of the research question?

The elements of the research question of each study should be identified, following the PICO framework (Table 9).

Table 9. Extracting the PICO elements of the research question

Element	Description
Population (P)	● When describing the study population, include important characteristics such as age, sex, ethnicity, geographic location, socioeconomic status, and study setting (hospital or community setting) ● Baseline clinical characteristics such as disease severity and comorbidities should also be noted.
Intervention/ Exposure (I/E)	● When describing the intervention or exposure, define the following as applicable: ○ Mode of administration ○ Dose and dosage forms ○ Frequency and duration of administration/exposure ○ Co-interventions ● Specify if the intervention or exposure is an alternative, adjunct, or sequential modality to the current standard of care.
Comparator/s (C)	● Comparators may include usual care, standard of care, another active treatment, or the same treatment given at a different dosage or modality.
Outcome (O)	● The relevant primary and secondary outcome/s measured in the study should be extracted. ● Clinical outcomes that are meaningful to patients, clinicians, and policy-makers are preferred (e.g. survival, need for hospitalization, quality of life, etc.) ● When appropriate, specify which tool was used to measure the outcome (e.g. depression measured using PHQ-9). Identify whether the outcomes were measured using standardized and validated tools. ● Surrogate endpoints (e.g. fasting blood sugar levels, serum creatinine levels) are acceptable only if there is clear epidemiological evidence linking them directly to the clinical outcome of interest. ● The timing of outcome assessment should also be specified (e.g. presence of depression at 3 years follow-up period)

6.4 How do we extract key components of the study methodology?

6.4.1 Study design

Identify the study design of each included RWE. The study designs of RWE that can be used for clinical evaluation of health technologies are varied, and may include pragmatic clinical trials, prospective or retrospective cohort studies, case-control studies and cross-sectional studies.

6.4.2 Inclusion and Exclusion Criteria

Extract the eligibility criteria, which includes the inclusion and exclusion criteria, used to select the study participants of each RWE. This information is crucial for evaluating the generalizability (external validity) of the RWE results.[2]

6.4.3 Data Source

It is important to identify the type of data source (whether primary or secondary source data), and the specific data source where the data of each RWE was obtained. The accessibility of data, whether it is public (free access), restricted (access conditional to approval), or private (accessible only to certain users), should also be noted.[3]

It is ideal to identify the data source for each of the main study variables (i.e. population variables, exposure variables, outcome variables) and describe whether the data was directly obtained from the data source, derived or coded. The processes used for derivation or coding, and how these processes were validated, should also be obtained. For instance, if cancer recurrence data was derived from an algorithm based on hospital resource use or treatment activity data, obtain the derivation process and reference the validation work behind the algorithm. The rationale for coding numerical variables into categories must be obtained. (e.g., prostate-specific antigen levels categorized into ≤10 ng/ml, 10-20 ng/ml, or >20 ng/ml based on risk stratification)

6.4.4 Data Collection Methods

It is important to extract information on the process of data collection and management in RWE. This includes details on data access and extraction, and data cleaning and validation (Table 10).

Table 10. Data Collection Methods[2]

Process	**Description**
1. Data access and extraction	● Identify the methods used to ensure transparency and quality of data collected. ○ Specify how data sources were accessed by the authors ○ Identify who collected the data (e.g., trained data managers, physicians, researchers), the frequency of data collection, and whether the data was collected manually or digitally. ● State the exact date when the data was extracted for the study (i.e. data cut-off date or specific access date) to allow critical appraisal of data maturity especially for time-dependent outcomes
2. Data cleaning and validation	● Describe the methods used for data cleaning and validation. ● Determine if the data was systematically curated for the specific study, using established curation manuals, dictionaries, or coding systems.

	• Report on quality control measures, including a description of how outliers and missing values were assessed and handled during the data collection process.

6.4.5 Data Analysis Methods

Identify the main data analysis methods used in the RWE. Identify how the data was described, including the measures of central tendency (mean or median), measures of dispersion (standard deviation or interquartile range). If available, extract details on the frequency distribution of the data. This information will give readers a clear understanding of how the data is distributed and summarized in the RWE.

Describe the type of statistical analysis used. In RWE, it is particularly important to identify if any statistical method was used to account for confounders (e.g. propensity score matching, inverse probability weighting, covariate adjustment using regression analysis, instrumental variable analysis).[4] Specify whether the analyses conducted are pre-planned or post-hoc. It is also relevant to obtain the pre-planned sample size requirements and the statistical power of the study. Reporting these elements can bolster confidence in the results.[2]

6.5 How do we extract the results in RWE?

From each RWE, extract the results of the relevant outcomes identified in your research question. For dichotomous and time-to-event outcomes, it is recommended to extract the adjusted effect estimates (adjusted odds ratio, adjusted relative risk or adjusted hazard ratio) and its standard error or 95% confidence interval. The raw numerical data (i.e. number of events observed per study group) and the p-values may also be extracted for completeness of results; however, the adjusted effect estimates should be used in pooling results. For continuous outcomes, it is recommended to extract the partial regression coefficients derived from the regression model, aside from extracting the mean and standard deviation of the outcome. The use of adjusted effect estimates and partial regression coefficients allows statistical adjustments of confounding variables to mitigate confounding bias in RWE (see Chapter VII for more details). If these values are not reported, check if the full data set used in the RWE is available in the results or in the supplementary materials. The full data set may also be requested directly from the authors to allow computation of the adjusted effect estimates or partial regression coefficients (as a last resort)

For all outcomes, it is important to extract the mean or median follow-up time (from the index date to the event or censoring date). It is also important to identify what variables were included in the regression models, and the measures of effect and the respective confidence intervals for these included variables. The results of additional analyses, such as adjustments or imputations for missing data, subgroup analysis and sensitivity analysis should also be obtained.

References

1. Kang J, Cairns J. Protocol for data extraction: How real-world data have been used in the National Institute for Health and Care Excellence appraisals of cancer therapy. BMJ Open. 2022;12:e055985. doi:10.1136/bmjopen-2021-055985.
2. Castelo-Branco L, Pellat A, Martins-Branco D, Valachis A, Derksen JWG, Suijkerbuijk KPM, et al. ESMO Guidance for Reporting Oncology real-World evidence (GROW). Annals of Oncology. 2023;34(12):1097-1112. doi:10.1016/j.annonc.2023.10.001.
3. Pongiglione B, Torbica A, Blommestein H, de Groot S, Ciani O, Walker S, et al. Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal. International Journal of Technology Assessment in Health Care. 2021;37(e62):11–18. https://doi.org/10.1017/S0266462321000301.
4. Lin L, Ahn J, Bayani D, Chan K, Choiphel D, Isaranuwatchai W, et al. Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in ASIA. 2020. https://hiper.nus.edu.sg/realise-guidance/. Accessed February 17, 2024.

Chapter VII

How do we synthesize and analyze real-world evidence?

7.1 Introduction

7.2 How do we perform qualitative evidence synthesis of RWE?

7.3 How do we perform quantitative evidence synthesis of RWE?

7.3.1 What data should we use for the meta-analysis of RWE?

7.3.1.1.How do we perform meta-analysis of dichotomous outcomes?

7.3.1.2 How do we perform meta-analysis of time-to-event outcomes?

7.3.1.3 How do we perform meta-analysis of continuous outcomes?

7.3.2 How should we pool the treatment effect estimates in RWE?

7.3.3 How do we address heterogeneity in RWE?

7.3.4 Can we use indirect effect estimates (obtained through network meta-analysis) for RWE?

7.4 How do we assess the certainty of evidence in RWE?

Chapter Summary

RWE usually involves more diverse populations and broader inclusion criteria, and is prone to confounding bias and other biases. Appropriate statistical methods are necessary to address these issues. Qualitative synthesis involving RWE must emphasize contextual factors like healthcare systems and patient demographics, which are crucial in understanding the applicability and generalizability of findings in real-world settings. When performing meta-analysis, the adjusted effect estimates or partial regression coefficients from an adjusted regression model should be used. The Generic Inverse Variance (GIV) method should be used to obtain the pooled effect estimate. The random effects model is preferred over the fixed effects model due to the substantial heterogeneity inherent in RWE. Network meta-analysis is not suitable for RWE because of the heterogeneity of RWE. The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework is used for certainty of evidence assessment.

7.1 Introduction

Systematic reviews of real-world evidence (RWE) aim to summarize existing research, which includes quantitative and qualitative studies. RWE usually involves more diverse populations and broader inclusion criteria to reflect real-world conditions. RWE is also prone to confounding bias and other biases. Thus, appropriate statistical methods are necessary to address these issues.[1] Consequently, synthesizing and analyzing evidence from RWE requires reviewers to carefully consider these characteristics of RWE to avoid misleading findings.

7.2 How do we perform qualitative evidence synthesis of RWE?

Qualitative or narrative synthesis is a method used in systematic reviews to summarize and integrate multiple studies using descriptive methods.[2] Unlike meta-analysis which aims to pool estimates, narrative synthesis explores patterns and relationships across studies, generating a coherent understanding of the evidence.[3] Both qualitative and quantitative synthesis should ideally be performed in systematic reviews of RWE. However, instances exist when only qualitative synthesis is feasible, such as when substantial heterogeneity in exposures, outcomes, or study designs exist, or when there is only 1 RWE relevant to the systematic review.

Narrative synthesis involving RWE must emphasize contextual factors like healthcare systems and patient demographics, which are crucial in understanding the applicability and generalizability of findings in real-world settings. Different methods such as textual descriptions, summary tables, and visual plots can be utilized to present the qualitative synthesis of RWE. Textual descriptions are used to summarize important findings and synthesize data into cohesive summaries. Summary tables are used to compare study characteristics, outcomes, and significant findings across the included RWE. Visual plots such as forest plots, bubble plots, and albatross plots help to illustrate relationships, trends, and evidence clusters. The Cochrane Handbook for Systematic Reviews of Interventions offers detailed guidance on how to effectively employ these methods.[4]

7.3 How do we perform quantitative evidence synthesis of RWE?

Quantitative synthesis involves methods that combine numerical data from multiple studies. The most commonly used method is meta-analysis, which uses statistical techniques to combine study results to provide a weighted pooled estimate of effect. Meta-analysis also allows for a quantitative examination of heterogeneity among included studies.[5] The following sections detail the unique considerations when conducting a meta-analysis of RWE.

7.3.1 What data should we use for the meta-analysis of RWE?

When performing a meta-analysis of RWE, it is essential that the combined effect estimate is adjusted for confounders. Thus, raw, unadjusted data (e.g., frequencies, percentages, means, standard deviations) from the included RWE should not be used in the meta-analysis of RWE because these values are likely to be biased. Instead, reviewers should use the adjusted estimates of treatment effects (e.g., adjusted odds ratios, adjusted relative risk) since these account for observed confounding and provide a more accurate measure of the effect.[4]

There are certain study designs in RWE where confounders may seemingly have been accounted for, such as pragmatic trials or observational studies that utilize propensity score matching. Pragmatic trials usually involve randomization; however, the uncontrolled environment in a pragmatic trial may still introduce bias and unknown confounding factors.[6] Similarly, while observational studies employing propensity score matching mimic randomized trials by balancing observed confounders between exposure groups, imbalance can still occur due to imperfect matching and from confounders not included in the propensity score model.[7] Therefore, even in these scenarios, the use of adjusted estimates in the meta-analysis is still the recommended approach.

7.3.1.1.How do we perform meta-analysis of dichotomous outcomes?

When performing meta-analysis of dichotomous outcomes, it is recommended to use treatment effect estimates from regression models that control for confounding. Specifically, the adjusted odds ratio (aOR) or adjusted relative risk (aRR), along with its corresponding standard error, should be obtained from each included study. If the standard error of the effect measure is not available, the 95% confidence interval of the adjusted estimate can be used instead. Afterwards, the adjusted effect measure and its standard error or 95% confidence interval estimate should be log-transformed. This transformation can be performed using tools like the RevMan Web calculator (see Figure 3). A detailed guide on using the RevMan Web calculator is available online.[8]

The log transformation is essential to convert the effect measures into a suitable form for meta-analysis. This is done for the following reasons: 1) to achieve symmetry and normality in effect sizes, 2) to make the effects additive rather than multiplicative, and 3) to increase the stability of the variances, resulting in more reliable pooled estimates.[4]

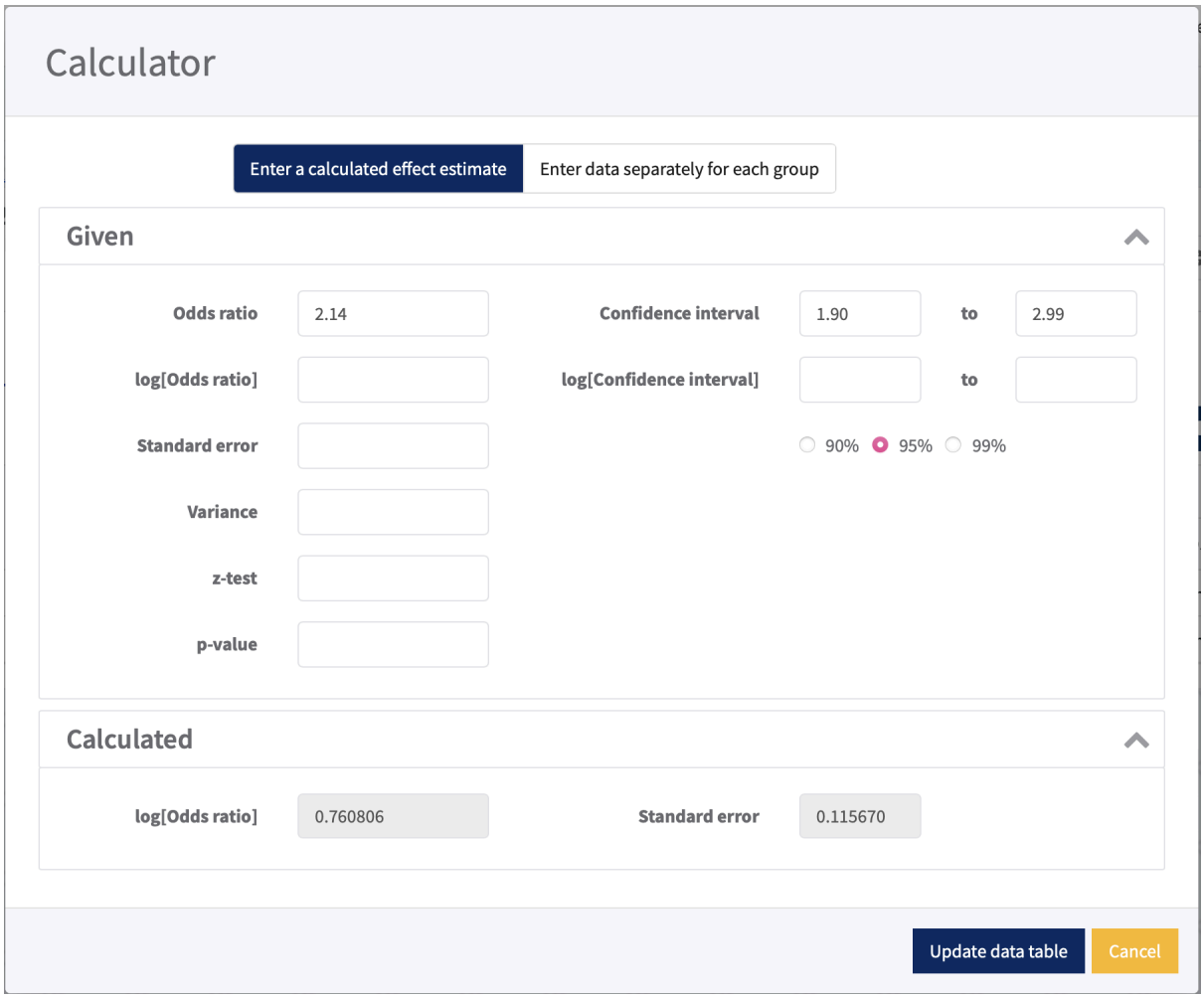

Figure 3. RevMan Web (Version 7.14.0) calculator for computing the log-transformation of the odds ratio and its standard error. In this figure, the RevMan Web calculator is used to convert OR=2.14 (95% CI: 1.90 to 2.99) to its log-transformed value of 0.76 with a standard error of 0.12.

7.3.1.2 How do we perform meta-analysis of time-to-event outcomes?

When performing meta-analysis of time-to-event outcomes, it is recommended to use the adjusted hazard ratio (aHR), which accounts for observed confounders, rather than the crude HR. The adjusted HR should be obtained along with its standard error or 95% confidence interval estimate.[4] Similar to dichotomous outcomes, the HR and its standard error or 95% confidence interval estimate should then be log-transformed. This transformation can also be performed using tools like the RevMan Web calculator (see Figure 4).

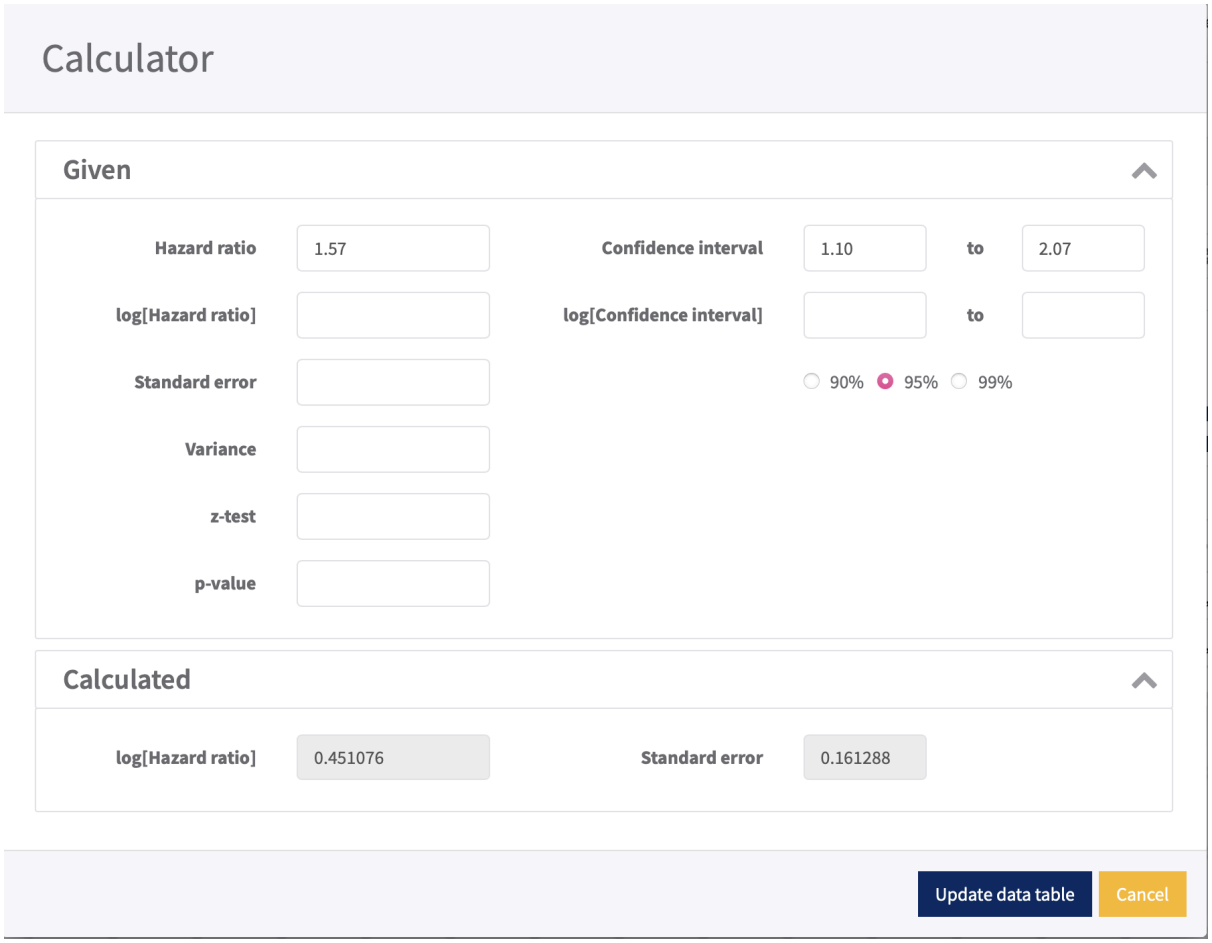

Figure 4. RevMan Web (Version 7.14.0) calculator for computing the log-transformation of the hazard ratio and its standard error. In this figure, the RevMan Web calculator is used to convert HR=1.57 (95% CI: 1.10 to 2.07) to its log-transformed value of 0.45 with a standard error of 0.16.)

7.3.1.3 How do we perform meta-analysis of continuous outcomes?

When performing meta-analysis of continuous outcomes, it is recommended to use the partial regression coefficients from an adjusted regression model instead of simply using the mean and standard deviation of the outcome variable. In RWE where a linear regression model was used, the partial regression coefficient can be interpreted as the mean difference (MD) between the exposure groups while controlling for confounders.[5] In addition to the partial regression coefficients, its standard error or 95% confidence interval should also be obtained.

For studies that use regression models other than linear regression models (e.g., multilevel models, spline regression models, robust regression models), it is necessary to perform additional data conversion to standardize effect sizes to a common metric. This ensures the comparability of the effect sizes across studies. It is recommended to convert partial regression coefficients from different regression models to standardized mean differences (SMDs) to enable uniform comparison across studies.[5]

However, converting partial regression coefficients from different regression models to SMDs can be complex and may require additional information beyond the regression coefficient and standard error. For example, converting the partial regression coefficients from multilevel models may require the intraclass correlation coefficient (ICC) and other relevant information. Therefore, it is advisable to consult a statistician for accurate conversion to SMDs.[9]

7.3.2 How should we pool the treatment effect estimates in RWE?

Since adjusted estimates from regression models are used in the meta-analysis of RWE, the Generic Inverse Variance (GIV) method should be used to obtain the pooled effect estimate. The GIV method is appropriate for continuous, dichotomous, and time-to-event outcomes because it allows for the combination of various types of effect measures (e.g., mean differences, odds ratios, hazard ratios) and accounts for the precision of each study's estimate by weighting them accordingly.[4]

The random effects model is preferred over the fixed effects model due to the substantial heterogeneity inherent in RWE.[5] RWE typically involves diverse study designs, populations, and exposures. The random effects model acknowledges and accommodates this heterogeneity by assuming that actual treatment effects vary across studies, yielding wider confidence intervals and more conservative estimates.[4]

7.3.3 How do we address heterogeneity in RWE?

As is common practice in meta-analysis, statistical heterogeneity should be quantified using methods such as the inconsistency statistic (I^2) and Cochran's Q test (i.e., chi-square test) [5]. A small p-value in Cochran's Q test (p-value <0.10) indicates significant statistical heterogeneity. The I^2 statistic represents the percentage of variability in effect estimates due to heterogeneity rather than chance. I^2 values of 50% or greater indicate significant heterogeneity[4]

Aside from reporting the above-mentioned statistics, subgroup analyses should also be employed to investigate sources of heterogeneity. Sensitivity analyses may also be performed to assess the robustness of findings by excluding studies with high risk of bias, varying analytical methods, and poorly controlled confounding. [10] Furthermore, if the number of included studies is sufficient, meta-regression may be conducted to adjust for confounding factors and identify possible sources of heterogeneity. [5]

7.3.4 Can we use indirect effect estimates (obtained through network meta-analysis) for RWE?

Network meta-analysis (NMA), which uses data from both direct and indirect evidence, is not suitable for RWE because of challenges related to methodological and data heterogeneity of RWE.[11] RWE often does not have direct comparisons between treatments, which makes it difficult to conduct reliable NMA. Moreover, NMA relies on the assumptions of transitivity (i.e., the concept that indirect comparisons are valid because the contributing studies are sufficiently similar) and consistency (i.e., the agreement between direct and indirect evidence). These assumptions are difficult to fulfill in the context of RWE due to their inherent variability.

7.4 How do we assess the certainty of evidence in RWE?

One of the most widely used tools for assessing the certainty of evidence from evidence synthesis is the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework. GRADE provides a systematic method for rating the certainty of evidence in systematic reviews and other evidence syntheses. It evaluates the level of certainty based on five domains (i.e., risk of bias, inconsistency, indirectness, imprecision, and publication bias) and categorizes the certainty of evidence into four levels (i.e., high, moderate, low, and very low).[12] GRADE can be applied to both RCTs and observational studies.

Due to the use of randomization and controlled environments, the rating of evidence for RCTs typically starts at a high level of certainty, but it can be downgraded based on issues identified in the five domains. The rating for observational studies, due to the lack of randomization, starts at a low level of certainty but can be upgraded depending on the methodological rigor of the included studies. A comprehensive guide to conducting evidence assessment using GRADE has been published previously.[12]

RWE is often observational and non-randomized, with common issues including confounding bias and reliance on secondary data sources with inherent limitations such as incomplete or missing data. [13] Thus, the GRADE assessment for RWE starts at a low level of certainty. This can be further downgraded based on issues identified in the five domains.

Despite the methodologic challenges in RWE, there are circumstances where the certainty of evidence can be upgraded. This can occur when studies exhibit large effect sizes, dose-response gradients, and the presence of plausible confounders that increases the certainty of effect.[12] However, it is important to note that once the rating has been downgraded to a very low level of certainty for any reason, the rating cannot be upgraded any further. Table 11 presents the factors for determining the certainty of evidence according to the GRADE Handbook.[12]

Table 11. Factors that affect the GRADE rating of the certainty of evidence

Factor	Consequence to the GRADE rating
Reasons for downgrading the level of certainty	
Limitations in study design or execution (risk of bias)	↓ 1 or 2 levels
Inconsistency of results	↓ 1 or 2 levels
Indirectness of evidence	↓ 1 or 2 levels
Imprecision	↓ 1 or 2 levels
Publication bias	↓ 1 or 2 levels
Reasons for upgrading the level of certainty	
Large magnitude of effect	↑ 1 or 2 levels
All plausible confounding would reduce the demonstrated effect or increase the effect if no effect was observed	↑ 1 level
Dose-response gradient	↑ 1 level

SOURCE: Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations: The GRADE Working Group; 2013. Available from: https://guidelinedevelopment.org/handbook.

References

1. Liu F, Panagiotakos D. Real-world data: A brief review of the methods, applications, challenges and opportunities. BMC Medical Research Methodology. 2022;22(1):287. doi:10.1186/s12874-022-01722-w
2. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme (Version 1). Lancaster. 2006.
3. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology. 2008;8:45. doi:10.1186/1471-2288-8-45
4. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. 2019.
5. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. 2011.
6. Ford I, Norrie J. Pragmatic Trials. The New England Journal of Medicine. 2016;375(5):454-463. doi:10.1056/NEJMra1510059.
7. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research. 2011;46(3):399-424. doi:10.1080/00273171.2011.570096.
8. Cochrane Training. (n.d.). RevMan Web Calculator. Retrieved from https://documentation.cochrane.org/revman-kb/calculator-95420946.html
9. Wilson DB. Regression coefficient from an HLM/Mixed Effects Model (Clustered Data). 2023. Retrieved from https://www.campbellcollaboration.org/escalc/d-regression-hlm.php.

10. Bowrin K, Briere JB, Levy P, Toumi M, Millier A. Use of real-world evidence in meta-analyses and cost-effectiveness models. Journal of Medical Economics. 2020;23(10):1053-1060. doi:10.1080/13696998.2020.1740518.
11. Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, et al. GetReal in network meta-analysis: A review of the methodology. Research Synthesis Methods. 2016;7(3):236-263. doi:10.1002/jrsm.1187.
12. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendations: The GRADE Working Group. 2013. Retrieved from https://guidelinedevelopment.org/handbook.
13. Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR Real-World Data Task Force report. Value in Health. 2007; 10(5):326-335. doi:10.1111/j.1524-4733.2007.00217.x.

Annexes

Annex 1. Methodologic filters (obtained from MEDLINE via PubMed)

Observational studies[11]	
1	epidemiologic studies [Mesh:noexp]
2	Cohort Studies[Mesh]
3	case control studies[MeSH Terms]
4	case control[Text Word]
5	cohort[tw] AND (study[tw] OR studies [tw])
6	cohort analy*[Text Word]
7	follow up [tw] AND (study[tw] OR studies[tw])
8	observational[tw] AND (study[tw] OR studies[tw])
9	longitudinal[Text Word]
10	retrospective[Text Word]
11	cross sectional[Text Word]
12	cross-sectional studies[MeSH Terms]
13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12

Pragmatic trials (Sensitivity-maximizing search)[12]	
1	((pragmatic*[tw] OR naturalistic[tw] OR real world[tw] OR real life[tw] OR unblinded[tw] OR unmasked[tw] OR cluster[tw] OR step* wedge*[tw] OR point of care[tw] OR factorial[tw] OR switchback[tw] OR switch back[tw] OR phase 4[tw] OR phase IV[tw]) AND (study[tw] OR trial[tw])) OR (practical trial[tw] OR effectiveness trial[tw] OR ((cluster*[tw] OR communit*[tw]) AND randomi*[tw]))
2	(general practice*[Text Word] OR primary care[Text Word] OR registry based[Text Word] OR health record*[Text Word] OR medical record*[Text Word] OR EHR[Text Word] OR EMR[Text Word] OR administrative data[Text Word] OR routinely collected data[Text Word] OR (communit*[Text Word] AND intervention*[Text Word]) OR quality improvement[Text Word] OR implementation[Text Word] OR decision support[Text Word] OR health service*[Text Word] OR health system*[Text Word] OR comparative effectiveness[Text Word] OR CER[Text Word] OR usual care[Text Word] OR evidence based[Text Word] OR practice guideline*[Text Word] OR (guideline*[Text Word] AND recommend*[Text Word]) OR knowledge translation[Text Word] OR health technology assessment[Text Word] OR HTA[Text Word] OR cost effectiveness[Text Word] OR process evaluation[Text Word] OR economic evaluation[Text Word] OR patient oriented[Text Word])
3	randomized controlled trial[pt] OR ((comparative effectiveness OR randomi*ed) AND (trial[Title] OR study[Title]))
4	(comment on[Title] OR phase 1[Title] OR phase I[Title] OR phase 2[Title] OR phase II[Title] OR non-randomi*ed[Title] OR quasi-randomi*ed[Title] OR pseudo-randomi*ed[Title]) OR (clinical trial, phase I[pt] OR clinical trial, phase II[pt] OR systematic review[pt] OR meta-analysis[pt] OR review[pt] OR editorial[pt])
5	pragmatic clinical trial[Publication Type]
6	((#1 OR #2) AND (#3 NOT #4)) OR #5
7	Animals[MeSH] NOT Humans[MeSH]
8	#6 NOT #7

Other studies[13]	
	((register* OR registr* OR database*)[Title/Abstract] OR (register* OR registr* OR database*)[MeSH Terms])
	(RWE[Title/Abstract]) OR (real world[Title/Abstract])) OR ("real practice"[Title/Abstract:~2])

Exclusion filter[14][a]	
1	Editorial[pt]
2	Letter[pt]
3	Randomized Controlled Trial[pt] OR Clinical Trial, Phase I[pt] OR Clinical Trial, Phase II[pt] OR Clinical Trial, Phase III[pt] OR Clinical Trial, Phase IV"[pt] OR Controlled Clinical Trial"[pt]
4	Comment[pt]
5	Case Reports[pt] OR classical article[pt] OR clinical conference[pt] OR collected works[pt]
6	congresses[pt] OR consensus development conference[pt] OR directory[pt] OR duplicate publication[pt] OR ephemera[pt]
7	guideline[pt] OR practice guideline[pt] OR historical article[pt] OR lectures[pt]
8	legal cases[pt] OR "legislation[pt]
9	news[pt] OR "newspaper article[pt] OR "patient education handout[pt] OR "personal narratives[pt]OR "pictorial works[pt]
10	video audio media[pt]OR "webcasts[pt]
11	Clinical Trials as Topic[Mesh] OR double-blind[All] OR placebo-controlled[All]
12	pilot study[All] OR pilot projects[Mesh]
13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12

[a] This is an example of an exclusion filter which you can use when there are too many articles that were retrieved. You may modify this example depending on the types of studies that are appropriate to include in your review.

Annex 2. Checklist of items for reporting RWE from pragmatic trials.[1]

Item	Section	Description
1	Title and abstract	Identify how participants were allocated to interventions (e.g., "random allocation," "randomized," or "randomly assigned") May or may not identify the study as a cluster randomized trial or a pragmatic clinical trial, as well as randomization scheme (e.g., parallel, stepped-wedge, adaptive)
Introduction		
2	Background	Explain scientific background and rationale Describe the health or health service problem that the intervention is intended to address and other interventions that may commonly be aimed at this problem
3	Objectives	Describe specific study objectives and hypotheses The target population, time horizon, setting should be mentioned
Methods		
4	Trial design	Describe the pragmatic aspects of the trial design: decisions related to the real-world healthcare setting, logistical considerations and clinical workflow, and service delivery. Explain the design, such as cluster randomization, stepped-wedge. Indicate if applicable whether this is a population-based study. If possible, include a schematic representation of the study design. For cluster randomized trials, define the clusters and describe how the design features apply to the clusters. For stepped-wedge, cluster randomized trials, define the timing and randomization of crossover from the control to the intervention. Describe important changes to the methods after the trial started, and include reasons
5	Stakeholder engagement	Eligibility criteria should be explicitly framed to show the degree to which they involved relevant stakeholders, such as typical participants and/or, where applicable, typical providers (e.g., nurses), institutions (e.g., hospitals), communities (e.g., localities or towns) and settings of care (e.g., different healthcare financing systems);
6	Participants	Describe eligibility criteria for participants; settings and locations where the data were collected Explain method of participant recruitment and attributes of the healthcare setting/system where data was collected
7	Interventions	Describe precise details of the interventions intended for each group and how and when they were actually administered; for cluster randomized trials, indicate if interventions were applied at the cluster level, individual participant level, or both Describe extra resources added to (or resources removed from) usual settings in order to implement intervention. Indicate if efforts were made to standardize the intervention or if the intervention and its delivery were allowed to vary between participants, practitioners, or study sites When relevant, include details on the experience and training (e.g., frequency, intensity) of those who delivered the intervention

		Describe the comparator in similar detail to the intervention
8	Outcomes	Clearly define primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors) Explain why the chosen outcomes and, when relevant, the length of follow-up are considered important to those who will use the results of the trial For cluster randomized trials, indicate whether the outcome measures apply to the cluster level, individual participant level, or both.
9	Sample size	How sample size was determined; explanation of any interim analyses and stopping rules when applicable If calculated using the smallest difference considered important by the target decision maker audience (the minimally important difference) then report where this difference was obtained For cluster randomized trials, describe the number of clusters and the cluster size, including whether equal or unequal cluster sizes are assumed. Indicate the intraclass (intracluster) correlation coefficient, as well as an indication of its uncertainty.
10	Randomization - sequence generation	Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification) Describe the type of randomization (e.g., individual, cluster, non-randomized). For cluster randomized trials, explain if stratification or matching was used.
11	Randomization - allocation concealment	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned For cluster randomized trials, specify that allocation was based on clusters. Indicate whether allocation concealment was at the cluster level, individual participant level, or both.
12	Randomization - implementation	Who generated the allocation sequence, who enrolled participants (or clusters), and who assigned participants to their groups For cluster randomized trials, describe how individual participants were included in the clusters, such as by random sampling or inclusion of all individuals identified as eligible.
13	Blinding (masking)	Whether participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment If blinding was not done, or was not possible, explain why
14	Statistical methods	Statistical methods used to compare groups for primary outcomes; methods for additional analyses, such as subgroup analyses and adjusted analyses For cluster randomized trials, indicate how clustering was considered.

15	Human subjects protection	Describe approval by an ethics committee (e.g., an institutional review board) as well as any other oversight bodies from which approvals were obtained. If the pragmatic trial involved a regulated product, indicate whether it was conducted under an investigational new drug (IND) authorization or its equivalent. Describe the method of authorization used for the use of protected health information, standards for data security, approach used for data monitoring and, if applicable, the existence of a data monitoring committee. For cluster randomized trials, indicate the nature of engagement with cluster representatives (e.g., discussion, consent) and whether consent was obtained from individual cluster members.
16	Use of data from EHRs or clinical and administrative information systems	If the source of data was from a clinical or billing database instead of one created primarily for research, describe: · Nature of the data source and data · Steps used in gaining permission to use the data · How the population of interest was identified (i.e., development of phenotype definitions, use of ICD-10 codes) · Any specific standards, data elements, or controlled vocabularies used, and provide details of strategies for translating across coding systems where applicable · Each clinical phenotype (i.e., EHR-based condition definition) used · Process for linking data from different sources, including EHRs, ancillary systems, administrative and billing systems, and external sources · Data management activities during the study, including a description of different data sources or processes used at different sites · Plan for archiving or sharing the data after the study, including specific definitions for clinical phenotypes and specifications for the coding system (name and version) for any coded data
Results		
17	Participant flow	Add flow diagram of participants and/or clusters through each stage—specifically, for each group, report the numbers of participants and/or clusters randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome; describe deviations from planned study protocol, together with reasons List number of participants or units approached to take part in the trial, the number which were eligible, and reasons for non-participation should be reported
18	Recruitment	List dates defining the periods of recruitment and follow-up Explain why the trial ended or was stopped
19	Baseline data	Add table showing baseline demographic and clinical characteristics of each group (and cluster, if applicable)
20	Unanticipated changes in care within study arms	Describe any unanticipated changes in care that occurred in the study arms that could affect the interpretation of the study; any intervention contamination and adjustments made to the analysis to accommodate contamination
21	Numbers analyzed	List number of participants or clusters (i.e., denominator) in each group included in each analysis and whether analysis was by “intention-to-treat”; state the results in absolute numbers when feasible (e.g., 10/20, not 50%)

22	Outcomes and estimation	For each primary and secondary outcome, summarize results for each group and the estimated effect size and its precision (e.g., 95% CI). For binary outcomes, give both absolute and relative effect sizes. For cluster randomized trials, provide results at the individual or cluster level, as applicable, and give the intraclass (intracluster) correlation coefficient for each primary outcome.
23	Ancillary analyses	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating which are prespecified and which are exploratory
24	Adverse events	Describe all important adverse events or side effects in each intervention group
Discussion		
25	Interpretation	Interpret the results while taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes Describe study relevance to decision-makers, as a defining component of a pragmatic trial is that it is intended to inform decision-makers about benefits, burdens, and risks of an intervention
26	Generalizability	Discuss generalizability (external validity) of the trial findings Describe key aspects of the setting which determined the trial results, possible differences in other settings where clinical traditions, health service organization, staffing, or resources may vary from those of the trial
27	Overall evidence	Add general interpretation of the results in the context of current evidence

[1] Based on the CONSORT extension statement for pragmatic trials and NIH Pragmatic Trials Collaborative criteria.

Annex 3. Checklist of items for reporting RWE from observational studies.

No	Item	For observational studies generating primary source data[1]	Observational studies using routinely collected health data (secondary source data)[2]	For observational studies using routinely collected health data specific to pharmacoepidemiologic research[3]
Title and abstract				
1	Title and abstract	1.a. Indicate the study's design with a commonly used term in the title or the abstract. 1.b. Provide in the abstract an informative and balanced summary of what was done and what was found.	*Additional items:* 1.c. The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. 1.d. If applicable, the geographical region and timeframe within which the study took place should be reported in the title or abstract. 1.e. If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	
Introduction				
2	Background and rationale	Explain the scientific background and rationale for the investigation being reported.		
3	Objectives	State specific objectives, including any prespecified hypotheses.		
Methods				
4	Study design	4.a. Include details of the specific study design (and its features) and report the use of multiple designs if used. 4.b. The use of a diagram(s) is recommended to illustrate key aspects of the study design(s), including exposure, washout, lag and observation periods, and covariate definitions as relevant.		
5	Setting	Describe the setting, locations, and relevant dates, including periods of recruitment, definition of time zero (i.e., date when subjects enter the study population) exposure, follow-up, and data collection.		
6	Participants	6.a Cohort study—give the eligibility criteria, and the sources and methods of selection, and follow-up of participants. For matched studies, give matching criteria and number of exposed and unexposed. 6.b Case-control study—give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and	6.a. The methods of study population selection (such as codes or algorithms used to identify participants) should be listed in detail. If this is not possible, an explanation should be provided. 6.b. Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.	Describe the study entry criteria and the order in which these criteria were applied to identify the study population. Specify whether only users with a specific indication were included and whether patients were allowed to enter the study population once or if multiple entries were permitted.

		controls. For matched studies, give matching criteria and the number of controls per case. 6.c Cross sectional study—give the eligibility criteria, and the sources and methods of selection of participants.	6.c. If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	
7	Variables	Clearly define all exposures of interest, outcomes, predictors, covariates (i.e.,. potential confounders and effect modifiers such as risk factors, comorbidities, comedications). Give diagnostic criteria, if applicable.	A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	7.a. Describe how the drug exposure definition was developed. 7.b. Specify the data sources from which drug exposure information for individuals was obtained. 7.c. Describe the time window(s) during which an individual is considered exposed to the drug(s). The rationale for selecting a particular time window should be provided. The extent of potential left truncation or left censoring should be specified. 7.d. Justify how events are attributed to current, prior, ever, or cumulative drug exposure. 7.e. When examining drug dose and risk attribution, describe how current, historical or time on therapy are considered. 7.f. Use of any comparator groups should be outlined and justified. 7.g. Outline the approach used to handle individuals with more than one relevant drug exposure during the study period.
8	Data sources / measurement	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	Specify data source name, data version, extraction date, sampling criteria,	Describe the healthcare system and mechanisms for generating the drug exposure records. Specify the care setting in which the drug(s) of interest was prescribed.
9	Bias	Describe any efforts to address potential sources of bias.		
10	Study size	Explain how the study size was arrived at.		
11	Quantitative variables	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.		

12	Statistical methods	12.a. Describe all statistical methods, including those used to control for confounding. 12.b. Describe any methods used to examine subgroups and interactions. 12.c. Explain how missing data were addressed. 12.d. Cohort study—if applicable, explain how loss to follow-up was addressed. Case-control study—if applicable, explain how matching of cases and controls was addressed. Cross sectional study—if applicable, describe analytical methods taking account of sampling strategy. 12.e. Describe any sensitivity analyses.		*Additional items:* 12.f. Describe and justify the use of multiple designs, design features, or analytical approaches.
13	Data access and management	Not applicable	13.a. Authors should describe the extent to which the investigators had access to the database population used to create the study population. 13.b. Authors should provide information on the data cleaning methods used in the study. 13.c. Specify procedures for securely receiving, quality checking, storing, backing up and preparing data. 13.d. Specify quality assurance procedures (e.g., double programming, assessing reliability of data by checking missing or miscoded data)	
14	Linkage	Not applicable	State whether the study included person level, institutional level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	
Results				
15	Participants	15.a. Report the numbers of individuals at each stage of the study (e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed). 15.b. Give reasons for non-participation at each stage. 15.c. Consider use of a flow diagram.	*Additional items:* 15.d. Describe in detail the selection of study participants including filtering based on data quality, data availability, and linkage. The selection of included individuals can be described in the text or by means of the study flow diagram.	
16	Descriptive data	16.a. Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders. 16.b. Indicate the number of participants with missing data for each variable of interest. 16.c. Cohort study—summarize follow-up time (e.g., average and total amount).		
17	Outcome data	Cohort study—report numbers of outcome events or summary measures over time. Case-control study—report numbers in each exposure category, or summary measures of exposure. Cross sectional study—report numbers of outcome events or summary measures.		

18	Main results	18.a. Give unadjusted estimates and, if applicable, confounder adjusted estimates and their precision (e.g., 95% confidence intervals). Make clear which confounders were adjusted for and why they were included. 18.b. Report category boundaries when continuous variables are categorized. 18.c. If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.		
19	Other analyses	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses.		
Discussion				
20	Key results	Summarize key results with reference to study objectives.		
21	Limitations	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Describe the degree to which the chosen database(s) adequately captures the drug exposure(s) of interest.
22	Interpretation	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.		*Additional item:* Discuss the potential for confounding by indication, contraindication or disease severity or selection bias (healthy adherer/sick stopper) as alternative explanations for the study findings when relevant.
23	Generalizability	Discuss the generalizability (external validity) of the study results.		
Other information				
24	Funding	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.		
25	Accessibility of protocol, raw data, and programming code (if any)	Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.		

[1] Based on STROBE (Strengthening the Reporting of OBservational studies in Epidemiology) and STaRT-RWE (Structured Template and Reporting Tool for Real-World Evidence)
[2] Based on RECORD (REporting of studies Conducted using Observational Routinely collected Data) and HARPER (HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence)
[3] Based on RECORD-PE (RECORD for pharmacoepidemiological research)

Annex 4. Data Extraction Form Template

Study ID	Name of Reviewer
GENERAL STUDY INFORMATION	
Title	
Authors	
Year of publication	Journal Name
Country/Region	DOI or URL
Time period covered by data observations	Funding agency
ELEMENTS OF THE RESEARCH QUESTION	
Population	
Intervention/Exposure	
Comparator	
Outcome	
STUDY METHODOLOGY	

Study Design

- □ Observational Study
  - O Prospective patient registries
  - O Retrospective patient registries
  - O Cross-sectional studies
  - O Prospective longitudinal cohort studies
  - O Retrospective longitudinal cohort studies
  - O Case-control studies
- □ Pragmatic clinical Trials
- □ Hybrid studies

Inclusion Criteria
Exclusion Criteria

Data source/s:

Study Variable	Data source	Type of data source (primary or secondary)	Accessibility of data source (public, restricted or private)

Data collection methods
1. Data access and extraction
- Who collected the data: _____
- Frequency of data collection: _____
- Type of data collection (i.e., manually or digitally collected): _____
- Date of study author's data extraction: _____

2. Data cleaning and validation
- Describe the methods used for data cleaning and validation: _____
- Describe quality control measures used (if any): _____

Data analysis methods

Planned sample size	Actual number of study participants

RESULTS

	Dichotomous outcomes	Adjusted effect estimates and standard error/95% CI	Exposure group n-		Control group n=	
			events	total	events	total
	Primary					
1						
	Secondary					
2						
3						

	Time to event outcomes	Adjusted effect estimates and standard error/95% CI	Exposure group n-		Control group n=	
			events	Person time	events	Person time
	Primary					

1						
	Secondary					
2						
3						

	Continuous outcomes	Partial regression coefficient	Regression model	Exposure group n-		Control group n=	
				mean	SD	mean	SD
	Primary						
1							
	Secondary						
2							
3							