Real-world data and real-world evidence in regulatory decision making

Report of the CIOMS Working Group XIII

The Council for International Organizations of Medical Sciences (CIOMS)

Geneva 2024

Real-world data and real-world evidence in regulatory decision making

Report of the CIOMS Working Group XIII

The Council for International Organizations of Medical Sciences (CIOMS)

Geneva 2024

ISBN: 978-929036106-0

Suggested citation: Real-world data and real-world evidence in regulatory decision making. CIOMS Working Group report. Geneva, Switzerland: Council for International Organizations of Medical Sciences (CIOMS), 2024.

https://doi.org/10.56759/kfxh6213. Licence: CC BY-NC-SA 4.0

CIOMS publications may be obtained directly from CIOMS through its publications emodule at https://cioms.ch/publications

Further information can be obtained from CIOMS, P.O. Box 2100, CH-1211 Geneva 2, Switzerland, www.cioms.ch e-mail: info@cioms.ch.

Design: Agence Gardeners, Annecy (France).

ACKNOWLEDGEMENTS

The Council for International Organizations of Medical Sciences (CIOMS) is thankful to the experts of this Working Group (WG) including government regulators, pharmaceutical professionals, academics, and health technology assessors, who contributed their valuable time and extensive experience to create this consensus report (see Appendix 3). During the course of the drafting, new members were invited to join the WG for their specific expertise.

The WG started work in March 2020 during the COVID-19 pandemic, and therefore began its meetings online, and having got used to this way of working, it continued mostly in the same way until the publication of this report. Warm thanks for everyone's dedication, flexibility and resourcefulness during the pandemic.

CIOMS gratefully acknowledges the Editorial Committee members who ensured the high quality of the report: Yoshiko Atsuta, Sean Hennessy, Juhaeri Juhaeri, and Jie Li. Special thanks go also to those WG members who served as chapter leads and co-leads, while valuing all input in the WG discussions, drafting, re-drafting and reviewing of the report. CIOMS also thanks Dinesh Mehta for his professional proofreading of the report.

The case studies in the Appendix 1 were submitted by WG members; however, in addition, we also welcomed contributions from the following WG members' colleagues and contacts from outside the WG:

- Nathalie Fretault, Novartis Pharma, France;
- Jurgen Kuball, The European Society for Blood and Marrow Transplantation (EBMT) / University Medical Center Utrecht, the Netherlands;
- Marcelo Pasquini, Center for International Blood and Marrow Transplant Research (CIBMTR), Medical College of Wisconsin, US; and
- Yoshiaki Uyama, Pharmaceuticals and Medical Devices Agency (PMDA), Japan.

CIOMS appreciates the comments of the many external reviewers (see Appendix 4) who participated in the public consultation of the draft report, that was hosted online on the CIOMS website during six weeks in 2023, and which was promoted by the WG's far-reaching networks. All comments and suggestions helped to inform the report and were carefully evaluated for their best implementation.

At CIOMS, Lembit Rägo, Sanna Hill, and Sue le Roux managed the project, with Kateriina Rannula providing technical support.

Lembit Rägo, MD, PhD
Secretary-General, CIOMS
Geneva, Switzerland, May 2024

TABLE OF CONTENTS

LIST OF TABLES, FIGURES AND BOXES

ABBREVIATIONS

AESI	adverse events of special interest
ALL	acute lymphoblastic leukaemia
Anvisa	Agência Nacional de Vigilância Sanitária [Brazilian Health Regulatory Agency]
BCMA	B-cell maturation antigen
BMI	body mass index
CADTH	Canadian Agency for Drugs and Technologies in Health
CAR-T	chimeric antigen receptor T-cell
CDM	common data model
CIOMS	Council for International Organizations of Medical Sciences
COVID-19	coronavirus disease
CRF	case report form
DARWIN EU	Data Analysis and Real World Interrogation Network
DLBCL	diffuse large B-cell lymphoma
EAP	Expanded access programme
eCRF	electronic case report form
EHDEN	European Health Data & Evidence Network
EHR	electronic health record
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
EUNetHTA	European Network for Health Technology Assessment
G-CSF	human granulocyte colony-stimulating factors
GDPR	General Data Protection Regulation (of the European Union)
GPP	Good Pharmacoepidemiology Practices
GPSP	Good Post Marketing Study Practice
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
GxP	Good 'insert activity' Practices. Guidances for Good Practices (general term that includes Clinical activity (GCP), Manufacturing activities (GMP), Pharmacovigilance (GVP), and others).
HARPER	HARmonized Protocol Template to Enhance Reproducibility
HCP	health care professional
HETE	hypothesis evaluating treatment effectiveness

HMA	Heads of Medicines Agencies
HRQoL	health-related quality of life
HTA	health technology assessment
ICD-10	International Classification of Diseases-10
ICD-11	International Classification of Diseases-11
ICER	incremental cost-effectiveness ratio
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICMRA	International Coalition of Medicines Regulatory Authorities
INESSS	Institut national d'excellence en santé et en services sociaux
IPTW	inverse probability of treatment weighting
IRB	institutional review board
ISPE	International Society for Pharmacoepidemiology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MA	marketing authorisation
MAH	marketing authorisation holder
MAR	missing at random
MHLW	Ministry of Health, Labour and Welfare (of Japan)
MID-NET®	Medical Information Database NETwork
MIHARI	Medical Information for Risk Assessment Initiative
MoCD	molybdenum cofactor deficiency
NDA	new drug application
NDB	National Claims Database (of Japan)
NDMA	N-nitrosodimethylamine
NICE	National Institute for Health and Care Excellence
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OS	observational study
PAES	post-authorisation efficacy study
PASS	post-authorisation safety study
PED	patient experience data
PIPEDA	Personal Information Protection and Electronic Documents Act
PMDA	Pharmaceuticals and Medical Devices Agency (of Japan)
PRO	patient-reported outcome
PROS	PIK3CA-related overgrowth spectrum
QoL	quality of life

RCT	randomised controlled trial
R&D	research and development
REC	research ethics committee
RMP	risk management plan
RNDS	Rede Nacional de Dados em Saúde [National Health Data Network of Brazil]
RWD	real-world data
RWE	real-world evidence
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SOC	standard of care
SRS	spontaneous reporting systems
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TGA	Therapeutic Goods Administration (of Australia)
UK	United Kingdom of Great Britain and Northern Ireland
UN	United Nations
US	United States of America
US FDA	US Food and Drug Administration
WHO	World Health Organization

PREFACE

Randomised controlled trials (RCTs) have long been considered the preferred approach for establishing the efficacy of new therapies. Most pre-approval RCTs are designed to demonstrate efficacy, or the ability of an intervention to produce a desired effect in expert hands under ideal circumstances. However, many research questions are difficult or impossible to address through RCTs, for example, where is no viable active comparator for an experimental treatment of a severe or life-threatening disorder. Further, some pre-approval trials are of limited duration, exclude high-risk populations, or have limited statistical power to detect rare but potentially serious adverse events in real-world patients. Recognising this, regulatory agencies have for many years accepted real-world evidence (RWE), which is derived from data collected outside of RCTs to fulfil post-approval safety requirements, and sometimes to demonstrate efficacy. In recent years, many medicines regulatory agencies have expressed increased willingness to consider RWE to support claims of efficacy or effectiveness as well as of safety. This increased willingness is changing the regulatory environment in which RWE is generated and used.

In the context of this changing regulatory environment, the data and methods used to generate RWE are changing as well. To assist those responsible for generating or interpreting RWE, the Council for International Organizations of Medical Sciences (CIOMS) has produced this consensus report on *Real-world data and real-world evidence in regulatory decision making*. The report introduces real-world data (RWD) and RWE (**Introduction**); describes RWE for decision making during the product lifecycle (**Chapter 1**); describes sources of RWD (**Chapter 2**); discusses RWE for regulatory use: key considerations (**Chapter 3**); discusses ethics and governance (**Chapter 4**); and provides conclusions and future directions (**Chapter 5**). While we are mindful of the rapid changes that affect RWD, methods for generating RWE, and the regulatory landscape in which RWE is applied, we hope that readers find this report useful.

EXECUTIVE SUMMARY

Introduction

Data from sources other than traditional randomised clinical trials are known as real-world data (RWD), and the evidence derived from the review and analysis of RWD is known as real-world evidence (RWE). RWD and RWE are used increasingly throughout the lifecycle of medicinal products to provide evidence about their effectiveness and safety. Recent regulatory guidance about RWE and approvals based on the use of RWE to demonstrate beneficial effects of products have created an urgency to develop generally accepted processes that promote trust in the evidence-generation process. This report describes the use of RWE for decision making in the lifecycle of medical products, describes RWD and data sources, discusses key scientific considerations in the generation of RWE, and discusses ethical and legal issues in using RWD.

Chapter 1: Real-world evidence for decision making during the product lifecycle

A number of stakeholders use RWE to support their decision making, including medical product regulators, health technology assessment (HTA) organisations, healthcare payers, patients, health care professionals (HPCs), and pharmaceutical companies[i]. Several medical product regulatory agencies have issued guidance on the key considerations for the use of RWE to support regulatory decisions. RWE can inform decisions at several points in a medical product's lifecycle. For RWE to support decision making, sponsors, regulators, and HTAs should implement a transparent process of planning for, reporting and evaluating RWE.

Chapter 2: Sources of real-world data

The scope of RWD is broad, including health care data and federated systems of health care data, spontaneous adverse event reporting systems, ad-hoc data collection, as well as emerging sources such as mobile devices and biosensors. An important challenge in using novel devices to generate RWD is the need to assure the validity of the resulting data.

Chapter 3: Real-world evidence for regulatory use: key considerations

RWD are often collected originally for reasons other than research. As a result, the fitness of specific RWD for specific research purposes needs to be assessed. The fitness of RWD depends on several factors including the research design that they will be used for. Chapter 3 outlines commonly used epidemiologic research designs and design elements, and discusses considerations for the statistical analysis of RWD. It also summarises current best practices regarding study registration, transparent reporting, documentation and responsible communication, as well as the increasing focus on improving the reproducibility of studies using RWD.

Chapter 4: Ethics and governance

Ethical and governance issues should be carefully considered when using RWD to generate RWE. These include not only privacy and data protection issues, but also informed consent as well as the efficacy-effectiveness gap between outcomes observed in RCTs (efficacy) and outcomes in real-world circumstances (effectiveness). One challenge is that while data protection laws protect the fundamental rights and interests of citizens in relation to the processing of their personal data,

[i] In the context of this report, pharmaceutical companies include manufacturers of prescription and non-prescription medicines.

the laws can be too restrictive to use RWD for evaluating the effectiveness of medicinal products on citizens' health.

Chapter 5: Conclusions and future directions

This report discusses the role of RWD/RWE in health-related regulatory decision making along the medicinal product's lifecycle and the needs of different stakeholders, available data sources, key scientific considerations, as well as the ethical and legal perspectives. More work remains to be done to globally harmonise practices and guidance for using RWD and RWE for regulatory decision making, thereby maximising the benefits they can bring to public health.

INTRODUCTION

To choose the best course of action, those making decisions about the approval, use, and reimbursement of medicinal products need to be able to weigh available evidence. Medicinal products are defined as substances or combinations of substances, including biological products, intended to treat, prevent or diagnose a disease, or to restore, correct or modify physiological functions by exerting a pharmacological, immunological or metabolic action.[1]

Regulators decide whether a medicinal product should be authorised for use, in which conditions or therapeutic indications, and for which patients and under which legal status (e.g., prescription or non-prescription). Healthcare payers decide whether an authorised medicinal product should be reimbursed, to whom, and at what price. Health care professionals (HCPs) decide whether they want to use a medicinal product, and for which patients. Finally, the patient — the ultimate decision maker in many circumstances — decides whether or not to use the product. Such decisions rely on evidence about the product's benefits and risks. To allow the most informed decisions, this evidence needs to be valid and unbiased, or if biased, the biases need to be understood and taken into account in the decision-making process.

How can valid evidence be obtained? For many years, randomised controlled trials (RCTs) have been considered to be the preferred source of evidence for evaluating the benefits of medicinal products, and are still widely viewed as the gold standard research design for such uses. Randomisation, the key feature of RCTs, provides some assurance that those randomised to different treatments are balanced, on average, with respect to baseline factors, whether measured or unmeasured, that could affect the study outcome. The likelihood of achieving such balance rises with the number of patients randomised.

RCTs usually test the efficacy of a new medicine against either a biologically inactive product, known as a placebo, or another medicine already authorised for the same indication. Participants are randomised at the start of the trial to one of the two or more treatment arms. Pre-specified data elements are often collected at fixed time points according to a detailed research protocol, which describes the analyses that will be performed. Beginning with the enactment of the 1962 Kefauver-Harris Drug Amendments to the US Food Drug and Cosmetic Act and analogous laws in other countries, RCTs became the norm for demonstrating efficacy.[2]

However, a number of important research questions are challenging or impossible to address with randomised designs. Further, a limitation of typical pre-approval clinical trials is that historically they tended to enrol participants who were not always representative of the population who would use the product once it is approved. This tendency has led to concerns about an *efficacy-effectiveness gap* between outcomes observed in RCTs (efficacy) and outcomes when the same medicine or intervention is used in real-world circumstances (effectiveness).[3]

Such concerns have prompted a change in the approach used to establish exclusion criteria, thus widening the trial population to make it more representative of the target patient population. Nevertheless, most pre-approval trials are still performed in relatively selected patient populations who are treated by highly selected clinical investigators under a strict research protocol. This has raised continuing questions about whether the findings are generalisable to patients, clinicians, and situations that are more commonly seen in the real world.

Another common limitation is that RCTs are designed with sufficient sample size to assess efficacy, and thus may not have enough statistical power to detect relatively infrequent but potentially serious

safety issues. To detect rare or unexpected events in the real-world setting and address them appropriately, studies utilising real-world data (RWD) are often needed.

All of this led to increasing use of RWD and real-world evidence (RWE), defined below, to inform regulatory and clinical decisions about medical products.

Definitions

Although various definitions of RWD have been proposed (see Table 1 for examples), there is currently no consensus definition.

Table 1. Some definitions of real-world data

Source: CIOMS Working Group XIII

Organisation	Definition of RWD
US Food and Drug Administration[4]	Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources. Examples of RWD include data derived from electronic health records, medical claims data, data from product or disease registries, and data gathered from other sources (such as digital health technologies) that can inform on health status.
European Medicines Agency[5]	Routinely collected data relating to a patient's health status or the delivery of health care from a variety of sources other than traditional clinical trials.
Joint International Society for Pharmacoepidemiology (ISPE) – International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Special Task Force on Real-World Evidence in Health Care Decision Making[6]	Data obtained outside the context of randomized controlled trials (RCTs) generated during routine clinical practice.
International Society for Pharmacoeconomics and Outcomes Research (ISPOR)[7]	Data relating to areas such as patient health status and/or healthcare delivery not collected in conventional RCTs. Examples of RWD are electronic health records (EHRs); wearables; medical claims data; surveys; and product, patient, and disease registries.
RAND Corporation[8]	Data collected during the routine delivery of care and its reimbursement. This type of data, referred to as real-world data, includes patient registries, electronic health records (EHRs), healthcare claims databases, and patient-generated data and is defined by its production outside of a research study.

Organisation	Definition of RWD
Innovative Medicines Initiative Get Real Project[9]	An umbrella term for data regarding the effects of health interventions (e.g. safety, effectiveness, resource use, etc.) that are not collected in the context of highly-controlled RCT's. Instead, RWD can either be primary research data collected in a manner which reflects how interventions would be used in routine clinical practice or secondary research data derived from routinely collected data. Data collected include, but are not limited to, clinical and economic outcomes, patient-reported outcomes (PRO) and health-related quality of life (HRQoL). RWD can be obtained from many sources including patient registries, electronic medical records, and claims databases.

RWE is evidence derived from the review and/or analysis of RWD.[10] These definitions of RWD are largely overlapping and similar in their essence. Given these and other existing definitions of this evolving field, the WG declined to put forth yet another definition.

Reasons for using real-world data and evidence

An important reason why decision makers may need to consider evidence from sources other than RCTs is that comparative trials of some interventions may not be possible because of ethical or logistical concerns. This may be the case if there is no viable active comparator for an experimental treatment of a severe or life-threatening disorder. An example of when a placebo arm in a trial was considered unethical occurred with avelumab for the treatment of Merkel cell carcinoma. When the trial commenced, there was no authorised medicine to act as a comparator, although another treatment was being developed by a separate company. The manufacturer of avelumab decided that a placebo arm would not be ethical, given that it could prevent a patient randomised to placebo from having the opportunity to receive an active treatment. The result was a single-arm trial that used a historical comparator group.[11]

Other reasons why a RCT may not be feasible is that some patients, particularly those with severe or life-threatening conditions, may be unwilling to enter placebo-controlled trials where there is only a 50% chance of getting the active drug. Finally, as mentioned above, decision makers may need to consider evidence from sources other than RCTs because efficacy as assessed in highly controlled trials may differ from real-world effectiveness, and, due to limited sample size, RCTs may not be suitable to evaluate safety events, especially the rare ones. At the time of medicine approval, questions about long-term safety may be unanswered.

For all these reasons, some have argued that decision makers should be more flexible in what evidence they rely upon, and use evidence both from randomised trials and other study designs to inform their conclusions.

In recent years, research designs have been refined and modified, and the boundary between RCTs and RWE has become more blurred. For example, a pragmatic trial in which patients are randomised and then followed up using routinely collected data has aspects of both a RCT and RWE. Thus, the range of study designs to answer a particular question now covers a wide spectrum of possibilities.

The great majority of RWD will come from products already on the market, because nearly all information on investigational medicinal products is collected in highly controlled manner. However, even investigational products can generate RWD (for example in the context of early access programmes), and their development can be complemented, accelerated and supported by relevant RWE.

RWD and RWE have been used for decades to characterise the adverse effects of medicinal products after their regulatory approval. The 21st Century Cures Act requires the US Food and Drug Administration (FDA) to establish a programme for evaluating RWE to support the approval of a new indication or evaluate extensions of an existing indication, but not for initial indications.[12]

Given that it is generally much less expensive to develop RWE than to perform RCTs to evaluate efficacy, pharmaceutical companies have a significant financial incentive to use RWE to support extensions of product indications. Further, the use of RWE to support initial marketing authorisations (MAs) has been introduced, most frequently in the context of a single-arm trial with a synthetic control arm, consisting of simulated patients or patients from outside the clinical trials of interest. In this context, the function of the synthetic control arm derived from RWD is to quantify the natural course of a disease or outcomes under the current standard of care (SOC).

However, as the actual and proposed use of RWD and RWE for supporting label claims for the effectiveness of medicinal products has increased, there has been significant debate as to whether and when such use is appropriate. For example, some argue that 'the replacement of randomised trials with non-randomised observational analyses is a false solution to the serious problem of ensuring that patients receive treatments that are both safe and effective',[13] even though approval decisions by regulatory agencies (including the US FDA[14]) have sometimes been based on non-randomised evidence even before the 21st Century Cures Act was passed.

CIOMS has developed this consensus report to inform discussions about the use of RWD and RWE for regulatory and health care decision making, including decisions to make a product available for use (authorisation), to cover the costs of its use (reimbursement), and to use a product for a particular patient (clinical use).

While using RWD is justified on ethical and logistical grounds, its use raises ethical and legal issues, which are also addressed in Chapter 4, including patient consent to the use of their data, privacy and data protection.

Regulatory potential of real-world evidence: controversies and challenges

RCTs have long been the mainstay for evaluating the efficacy of a medicinal product and are often a prerequisite for obtaining a licence to market a medicine. Randomisation reduces the possibility of imbalances among treatment groups, thus avoiding biased study results. However, RCTs often exclude large proportions of the general disease population and specific patient groups from trial participation including older adults, women, and patients with co-morbidities.[15] As Eichler and colleagues noted, restricting study populations 'increases the ability to detect a drug effect if it is there but reduces external validity. Progressive reduction of those uncertainties will need to be achieved by way of the use of [real-world] data from observational studies.'[16]

The uncertainties that Eichler and colleagues refer to concern the potential benefits and risks, as well as how a medicine will perform and be used in 'real life'. It is usual, at the time of authorisation of a medicine, for efficacy (the performance of an intervention under ideal and controlled circumstances)

to have been shown in the population studied, but its effectiveness (performance under real-world conditions) to be largely unknown, although hoped to be similar to its efficacy.[17]

The safety profile of a medicine is often less well known because of both the large study sizes needed to detect less common adverse effects, and the exclusion from clinical trials of people most likely to be at risk of harm, including older adults, children, pregnant and lactating females, and people with concomitant illnesses and/or on concomitant medication. Many adverse effects, especially rare ones, will be detected only once a medicine is used in real life in larger numbers and varieties of patients. For this reason, in many jurisdictions the unknowns about the safety profile will be researched after authorisation and, for that purpose specified in risk management plans (RMPs): documents that provide information on a medicine's safety profile, describe the activities of the marketing authorisation holder (MAH) to further characterise the safety profile after approval, and explain the measures to prevent or minimise the medicine's risks in patients. RMPs may also include mandated studies on aspects of efficacy.[18]

As mentioned, the utility of RWE is being increasingly recognised by regulatory bodies. The US 21st Century Cures Act of 2016 emphasises the use of RWE to support regulatory decision making, including approval of new indications for approved drugs. Based on this, in August 2023 the US FDA issued *Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products* to provide guidance on the use of RWE to support approval of a new indication for previously approved drugs.[19]

Similarly, in 2017 the European Medicines Agency (EMA) and Heads of Medicines Agencies (HMA) established a joint task force, later superseded by The Big data Steering Group, to describe the big data landscape from a regulatory perspective, and to identify how to optimally use big data in support of innovation and public health in the European Union (EU).[20]

In addition, EMA issued its guideline on registry-based studies in September 2021.[21] It focuses on studies using registries as a data source with a possible regulatory purpose.

To help harmonise efforts, in June 2022, the International Coalition of Medicines Regulatory Authorities (ICMRA), issued a statement on international collaboration to enable RWE for regulatory decision making. (See Appendix 2.)

Typically, RWD has been used to fulfil post-approval requirements and conduct long-term follow-up studies if there is remaining uncertainty about risks and duration of benefits at the time of approval. Increasingly though, RWD/RWE is applied to capture clinical outcomes in pragmatic and large simple trials, which enrol large number of participants and minimise the amount of information collected on each participant.

More recently, RWD is also used to collect information on natural history of disease to be used as external controls where the use of a randomised comparator arm is impractical or unethical, such as rare cancers or other unmet medical needs, or ultra-rare diseases where there are not enough patients to conduct adequately powered trials.

There is a growing number of examples of effective use of RWE to support and drive regulatory decisions, not only for label extensions, but also accelerated and full approvals.

However, the use of RWE for documenting the beneficial effects of medical products is not without controversy; debate about quality and hierarchy of the various research designs and data sources for clinical evidence continues. Conventional perspectives, combined with existing regulatory and ethical standards, and legal risks may deter the use of RWE where it could provide a valid source of evidence for beneficial effects.

Concerns about robustness and interpretability of RWE remain, due to the inherent bias and confounding in non-randomised studies, in addition to missing data; these concerns can be only partially addressed with design and analysis methods. Other technical issues provide challenges, such as lack of standardisation across different RWD sets, or the comparability of multiple data sources when using RWD for external controls for clinical trials.

In addition, the use of health care data can raise concerns about data privacy. Efforts are currently underway to streamline consistent approaches to generation and use of RWE.

However, especially in areas of unmet medical need such as rare disease treatments or urgent situations like the COVID-19 pandemic, it is increasingly being recognised that there may not be a large enough patient base, or enough time to gather evidence for approval considerations in the traditional way. In such circumstances, RWE can be of value in informing the benefit–risk balance in the target population.

With increasing availability and accessibility of RWD as well as evolving methods and analytical capabilities, the role of RWD in clinical development and regulatory decision making is likely to increase. Especially promising is the development of study designs that combine the benefits from RCT and RWD while minimising the limitations of each. As this is yet relatively uncharted territory, it is critical to seek early consultation with regulators on acceptability of RWE as part of the evidence for efficacy, safety, or both.

Although the application of RWE to answer remaining significant uncertainty about benefit–risk balance upon approval is more accepted, often some discussion on the value of RWE to meet post-marketing requirements is useful.

Target audience and aims

This report is in keeping with CIOMS's mission to advance public health through guidance on health research and policy including ethics, medicinal product development and safety. CIOMS is an international, non-governmental, non-profit organisation established jointly by the World Health Organization (WHO) and the United Nations Educational, Scientific, and Cultural Organisation (UNESCO) in 1949. CIOMS represents a substantial proportion of the biomedical scientific community through its member organisations, which include many of the biomedical disciplines, national academies of sciences, and medical research councils.

The intended audience for this report includes medicinal product regulators, healthcare payers, health care and medicinal products industries, researchers, bioethicists, patients and HCPs, who produce RWE or use it to inform regulatory, reimbursement, or clinical decisions. This report aims to describe the use of RWE for decision making, describe RWD and data sources, discuss key scientific considerations in the generation of RWE, and discuss ethical and legal issues in using RWD. While the main focus of this report is use of RWE to evaluate medicinal products, many of the considerations discussed in this report can also be applied to medical devices.

Scope and structure of this report

This report covers the relevant aspects pertaining to the use of RWE for approval, use, and reimbursement of medicinal products. The report consists of five chapters following this introduction:

- Chapter 1 addresses real-world evidence for decision making during the product lifecycle;
- Chapter 2 addresses sources of real-world data;
- Chapter 3 discusses real-world evidence for regulatory use: key considerations;
- Chapter 4 addresses ethics and governance;
- Chapter 5 provides conclusions and future directions.

Introduction – References

1 European Medicines Agency (EMA). Glossary of Regulatory Terms. (Website accessed 15 September 2022).
2 Greene JA, Podolsky SH. Reform, regulation, and pharmaceuticals—the Kefauver–Harris Amendments at 50. *The New England Journal of Medicine*. 2012;18;367(16):1481. *http://doi:* 10.1056/NEJMp1210007.
3 Templeton AJ, Booth CM, Tannock IF. Informing patients about expected outcomes: the efficacy-effectiveness gap. *Journal of Clinical Oncology*. 2020;20;38(15):1651-1654. https://doi:10.1200/JCO.19.02035.
4 US FDA. Real-world evidence. (Website accessed 16 May 2024).
5 Cave A, Kurz X, Arlett P. Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe. *Clin Pharmacol Ther*. 2019 Jul;106(1):36-39. https://doi.org/10.1002/cpt.1426.
6 Berger ML, Sox H, Willke RJ, Brixner DL, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf*. 2017 Sep;26(9):1033-1039. doi: 10.1002/pds.4297.
7 Modified from: The Professional Society for Health Economics and Outcomes Research (ISPOR). *About Real-World Evidence*. (Website accessed 15 September 2022).
8 Timbie JW, Kim AY, Concannon TW. Use of real-world evidence for regulatory approval and coverage of medical devices: a landscape assessment. *Value in Health*. 2021;1;24(12):1792-1798. DOI: 10.1016/j.jval.2021.07.003.
9 Modified from: Makady A, de Boer A, Hillege H, Klungel O, et al. (on behalf of GetReal Work Package 1). What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value in Health*. 2017 Jul-Aug;20(7):858-865. https://doi.org/10.1016/j.jval.2017.03.008.
10 Modified from: Zuidgeest MGP, Goetz I, Meinecke AK, Boateng D, et al. The GetReal Trial Tool: design, assess and discuss clinical drug trials in light of Real World Evidence generation. *J Clin Epidemiol*. 2022 Sep;149:244-253. https://doi.org/10.1016/j.jclinepi.2021.12.019.
11 European Medicines Agency (EMA): Assessment Report. Bavencio. EMA/496529/2017 (PDF accessed 15 September 2022).
12 Authenticated U.S. Government Information (GPO). Public Law 114-255. 114th Congress. To accelerate the discovery, development and delivery of 21st century cures and for other purposes. 2016. (PDF accessed 15 September 2022).
13 Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *The New EnglandJournal of Medicine*. 2020;382:674-678. http://doi: 10.1056/NEJMsb1901642.
14 Belson NA. FDA's Historical Use of "Real World Evidence." Food and Drug Law Institute. (Website accessed 15 October 2022).
15 Kennedy-Martin, T., Curtis, S., Faries, D. et al. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 16, 495 (2015). https://doi.org/10.1186/s13063-015-1023-4.
16 Eichler HG, Baird LG, Barker R, Bloechl-Daum B, Børlum-Kristensen F, Brown J, et al. From adaptive licensing to adaptive pathways: delivering a flexible life-span approach to bring new drugs to patients. *Clinical Pharmacology & Therapeutics*. 2015;97(3):234-246. https://doi.org/10.1002/cpt.59.
17 Templeton AJ, Booth CM, Tannock IF. Informing patients about expected outcomes: the efficacy-effectiveness gap. *Journal of Clinical Oncology*. 2020;20;38(15):1651-1654. https://doi:10.1200/JCO.19.02035.
18 The European Parliament and the Council of the European Union. Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council. (PDF accessed 15 October 2022).
19 U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Oncology Center of Excellence (OCE), August 2023, Real-World Data/Real-World Evidence (RWD/RWE). Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry. 88 FR 60215. (PDF accessed 11 March 2024).
20 European Medicines Agency (EMA). How we work: Big Data. (Website accessed 15 October 2022).
21 European Medicines Agency (EMA): Guideline on registry-based studies. EMA/426390/2021 (PDF accessed 12 April 2024).

CHAPTER 1.
REAL-WORLD EVIDENCE FOR DECISION MAKING DURING THE PRODUCT LIFECYCLE

Health-related RWE, which can be derived from RWD, can potentially be used for a broad range of purposes due to different decision-making infrastructures across healthcare systems worldwide. A wide variety of data are now being routinely collected across multiple disease areas and clinical settings. Ongoing efforts to structure data, standardise their quality, and ensure interoperability (the ability of two or more components or systems to exchange information and to use the information that has been exchanged[1]) will further increase the potential value of RWD and RWE, and their use by decision makers.

Multiple stakeholders within health systems globally are beginning to use RWD and RWE in different ways. In recent years, the development of medicinal products and diagnostics have involved innovative applications and increased the utility of RWD and RWE during different stages along the development lifecycle, as outlined in Section 1.4 below.

This chapter outlines some real-world examples along the product lifecycle, which starts with the discovery, and concludes with the end of the marketing phase, highlighting how different stakeholders have used RWD and RWE in decision making for medicinal products and diagnostics. First, in our discussion of evidentiary requirements, we outline stakeholders' roles and expectations. Then, we highlight differences in types of decisions for which the information is used.

Next, we discuss frameworks that stakeholders may use for the acceptance of RWE, including examples of specific frameworks from individual countries and how such frameworks can adapt to respond to evolving or urgent health needs of the population. We discuss the planning of global RWE generation, including relevant decision points in the product lifecycle, specific stakeholder evidentiary needs, and the importance of, and mechanisms for, cross-stakeholder interaction and collaboration. We present examples of RWE along the product lifecycle, describe potential routes to engage with regulators/HTA bodies, and we provide recommendations on how and when they should be considered.

1.1. Regulators, health technology assessors, payers and other stakeholders

A variety of stakeholders is involved in decision making in different jurisdictions. These stakeholders play specific roles in the decision-making process and thus may have different expectations and requirements concerning evidence standards during the product lifecycle, which consists of product introduction, growth, maturity, decline, and prescription to over-the-counter switch. Moreover, within any given health system, they may have divergent views on the potential role of RWE in informing decision making.[2]

This chapter considers roles in relation to RWE and decision making for the following stakeholders: regulators, HTA bodies, payers, clinicians, patients and pharmaceutical companies.[i]

1.1.1. Regulators

Regulatory bodies, such as the EMA and US FDA, assess the evidence to determine if a medicine's benefit outweighs its risk for specific indications. If this balance is positive, the medicine is authorised for entry into their respective market. The fundamental goal of structured benefit–risk assessment is to ensure that the benefits of the medicine outweigh the risks for the patients throughout its lifecycle. Continual assessment and monitoring of the benefit–risk balance necessitates the ability to evaluate different types of data from multiple sources.

Regulatory bodies such as EMA and US FDA use an 'effects table' for structured benefit–risk assessment. The effects table summarises favourable and unfavourable effects for the product under assessment and the comparator, which the regulators have taken into account. Uncertainties are also described. The structured benefit–risk assessment is also mentioned in the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidance,[3] and is a continuous process that includes consideration of the therapeutic context, including the disease or condition, the available therapies, the unmet medical need, and the outcomes of the main studies. The ultimate purpose of the effects table is to make clear and transparent the grounds on which a benefit–risk assessment is made. RWE is sometimes included in the effects table as well as data from RCTs.

1.1.2. Health technology assessors

CIOMS Working Group XIII defines HTA as 'a multi-disciplinary process to determine the relative value of an intervention developed to prevent, diagnose or treat medical conditions; promote health; provide rehabilitation; or organise healthcare delivery'. The intervention can be a test, device, medicine, vaccine, procedure, programme or system.

The role of HTA organisations is to promote an equitable, efficient health system that offers high quality care by assessing the value of the interventions, if adopted for use, and to make recommendations for its appropriate use. The value of a medical product may be assessed at different points in its lifecycle, using data from a variety of sources and involves a multi-disciplinary process.[4] A HTA may examine the intended and unintended consequences of using a new health technology compared to existing alternatives.

An initial value assessment will often consider not only clinical efficacy and safety, but also costs and economic implications, ethical, social, cultural, and legal issues, organisational and environmental aspects, as well as wider implications for the patient, relatives, caregivers, and the population.[5] Reassessment will often involve evaluation of comparative effectiveness data. Importantly, estimates of value may vary depending on the perspective taken, the stakeholders involved, and the decision context.[6] Ascertainment of value is generally based on integrating various types of information including patient and clinical expert opinion, clinical trial data, as well as scientific literature and data from the real-world care setting.

[i] In the context of this report, pharmaceutical companies include manufacturers of prescription and non-prescription medicines.

1.1.3. Payers

In healthcare, a payer is a person, organisation, or entity that pays for the care services provided by a HCP. It most often refers to government or private insurance companies, which provide customers with health plans that offer cost coverage and reimbursements for medical treatment and care services. Additional costs borne by patients and their families to access care can be a consideration in determining the value.

Globally, the role of payers is to determine the access to medicines based on reimbursement, budget and pricing. Depending on the local established healthcare system, different models exist such as single payer (e.g. Taiwan and the UK) or hybrid models (e.g. Australia), but the ultimate goal is to provide cost coverage and reimbursements for medical treatment and care services.

The decision to add a medicinal product into a health plan is mainly determined by the value of a medicine based on an unmet need, clinical evidence, cost-effectiveness, overall budget impact and willingness to pay. Approaches may vary across different countries and across payers within the same country. Moreover, negotiations between payers and pharmaceutical companies may lack in transparency, and patient access and physicians' prescribing practices may evolve following payers' determination of a product's value. More transparent planning and use of RWD would be beneficial for improved coverage decisions.

1.1.4. Patients and health care professionals

The ultimate stakeholders are, of course, patients and HCPs who consume and prescribe medicines to improve health and wellbeing. The essential goal of informed decision making is to promote treatments to individuals that benefit the most and in the safest possible manner.

Patients and care providers can play a major role in the RWE landscape. The incorporation of patients', clinicians' as well as other stakeholders' perspectives in the generation of evidence — from the elaboration of research questions to the collection of patient-centred outcomes — help to provide more relevant results for decision making.

Technologies, such as wearable devices, are now available to capture valid RWD from patients in real-world settings, contributing to RWE generation.[7,8]

1.1.5. Product developers and marketing authorisations holders

A MAH is a company or other legal entity that has been granted permission by a regulatory authority to market a medicine or a vaccine in a national or regional territory.[9] In some regions, MAHs are also responsible for medical devices including diagnostics. MAHs provide evidence to answer questions posed by other stakeholders. This data can come from a variety of sources including RWD.

MAHs are responsible for ensuring that they, and any parties working for them, comply with all relevant standards legislation and guidelines (e.g. "good 'insert activity' practices", or GxP). Compliance with these standards ensures the reliability and integrity of the data (pre- and post-marketing) and production processes that support the authorisation of medicines and their quality, safety and effectiveness once on the market.

1.2. Evidentiary requirements by regulators and health technology assessors

1.2.1. Regulatory frameworks and guidance for real-world evidence

The evidentiary requirements and submission process for regulatory approval and for HTA have similarities but also some important differences, which are reflected in the variation of acceptance and use of RWD and RWE for decision making, depending on the context.

In general, all the accumulated evidence will be appraised, with both clinical trial data and RWD being part of an information continuum. However, evidentiary requirements may vary depending on the stakeholders involved and the geographical context; this is because regulators, HTA organisations and payers in different jurisdictions may have different opinions on the value of RWD and RWE. Below, we describe a few available resources on frameworks and guidance.

Regulators are constantly working on providing requirements and recommendations to improve and structure the use of RWD in decision making. In the regulatory context, RWE has mainly been used to provide safety information. However, in recent years, an increasing number of submissions have included RWE to provide evidence of effectiveness.[10,11]

In December 2018, the US FDA published a Framework[12] for evaluating the potential use of RWE to support the approval of a new indication for a medicine already approved or to support or satisfy post-approval study requirements. The US FDA Framework proposes three key considerations to evaluate RWE: (1) whether the RWD are appropriate for the proposed use; (2) whether the study design used to generate RWE can provide adequate scientific evidence to answer or help answer the regulatory question; and (3) whether the study meets regulators' requirements, such as those concerning the quality of study monitoring and data collection.

In late 2021, the US FDA issued four draft RWD guidance documents for industry on aspects of RWD and RWE in regulatory decision making:

- *Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products* discusses use of electronic health records and claims databases, including recommendations on how to select appropriate RWD sources and to define and validate study variables;[13]
- *Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products* is guidance for using registries (populations defined by disease, condition or exposure, followed over time to evaluate specified outcomes) that collect data in a standardised manner for a population defined by a disease, condition, or exposure;[14]
- *Data Standards for Drug and Biological Product Submissions Containing Real-World Data* focuses on US FDA-supported data standards in medicine submissions with data derived from RWD to promote compliance with relevant legal requirements;[15]
- *Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drugs and Biological Products* provides the US FDA's current thinking regarding regulatory considerations for non-interventional studies involving RWD.[16]

In September 2022, the US FDA published a final guidance entitled *Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products: Guidance for Industry*.[17] In early 2023, a draft guidance was published on externally controlled trials.[18]

In Europe, EMA sometimes requests post-authorisation efficacy studies (PAES) to generate evidence needed for standard benefit–risk assessment, or at least to complement it. PAES are conducted to address scientific uncertainties identified by EU regulators on aspects of the evidence of benefits that should be, or can only be, addressed after authorisation. In 2016, EMA developed an associated scientific guidance to support MAH in the design of PAES.[19]

In 2015, EMA established the Patient Registry Initiative to explore ways of expanding the use of patient registries by supporting a systematic and standardised approach to their contribution to the benefit–risk evaluation of medicines. EMA has finalised in 2021 a guidance on the use of registry-based studies.[20]

Opportunities for improving the utilisation of RWD were recently analysed in the wider context of Big Data. The HMA–EMA Joint Big Data taskforce operated from 2017 until December 2019 and aimed to describe the big data landscape from a regulatory perspective to ensure the EU regulatory system has the capability and capacity to guide, analyse and interpret these data.[21]

Big data as discussed by the taskforce included data such as electronic health records (EHRs), registry data and claims data, pooled clinical trials data, datasets from spontaneously reported suspected adverse drug reaction reports, and genomics, proteomics, and metabolomics datasets. Big data was seen to complement clinical trials and offer major opportunities to improve the evidence upon which we take decisions on medicines. It was stressed that understanding the quality and representativeness of big data would allow regulators to select the optimal data set(s) to study an important question impacting the benefit–risk balance of a medicine.

The HMA–EMA Joint Big Data taskforce concluded with 10 priority recommendations,[22] several of which are relevant for the future use of RWD. The HMA-EMA Joint Big Data Steering Group was set up in 2020 to oversee the implementation of the recommendations from the taskforce report. In the current context of lack of specific guidance for the use of RWD and RWE in pre-approval setting, EMA encourages marketing authorisation applicants to approach the Agency early in setting up their evidence-generation plans.

In addition to these guidelines, ICH also has several guidelines that refer to the use of RWD for supporting benefit–risk assessment discussions, including using RWD in clinical trials (ICH E8 R1 and E6 R3) and, the guideline on general principles on pharmacoepidemiological studies that use RWD for safety assessment of medicines (ICH M14). However, there seems to be no overarching ICH guideline that refers to the various guidance that explain how RWD can be used to support clinical trials designs and medicine development.[23]

1.2.2. Health technology assessors and real-world evidence

In the context of HTA and decisions concerning reimbursement, data from real-world sources have been used to contextualise information to a specific regional health care setting; initiatives to generate RWE to fill gaps in evidence are increasing.[24] For example, the Commissioning through Evaluation programme in England enables new clinical and patient experience data (PED) to be collected for treatments that show promise but are not currently routinely funded due to significant uncertainties concerning clinical- or cost-effectiveness. The Australian government introduced a managed entry scheme as early as 2010 to gather evidence to resolve uncertainties for medicines to treat conditions of high and unmet clinical need.

Different regions around the world such as some Asian countries, Canada, and the UK are developing and publishing their own frameworks to guide the use, generation, reporting and appraisal of RWE for decision making.[25,26] In 2022, UK National Institute for Health and Care Excellence (NICE)

published its real-world evidence framework.[27] Health Canada and Canadian Agency for Drugs and Technologies in Health (CADTH) have established a RWE Steering Committee to optimise the use of RWE for regulatory and HTA decision making.

Many stakeholders are still learning how to optimise the integration of RWD and RWE into HTAs. There are examples where RWD and RWE have informed decision-making processes, but also examples where such data was insufficient to support a decision because, for example, the methodology for collecting and analysing the data was not considered appropriate or the quality of the data was not of an acceptable level.

Local and regional differences in HTA approaches to medicines value assessment present additional complexity for medicines manufacturers and developers. In the current environment, it is almost impossible for sponsors involved in new product commercialisation to have a common global evidence strategy targeting all stakeholders. Familiarity with local culture and historical experience with a country's HTA is needed to tailor evidence generation strategy and understand the expectations and uses of RWD and RWE locally. In a recent review of the use of RWE to inform NICE's cancer medicine appraisals from 2011 to 2018,[28] RWE was rarely rejected, but there was frequent criticism of the submitted RWE that was typically related to data sources and its relevance to inform the decision problem.

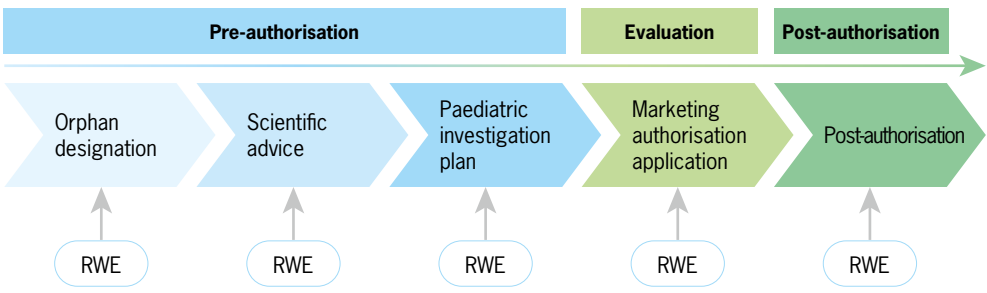
1.3. Planning for real-world evidence during product development

1.3.1. Relevant decision points in product lifecycle

Ideally, for each development programme, the sponsor should establish the evidence needed for regulatory approval, including RWE, at each of the different decision timepoints in the product lifecycle. While some evidence gaps might need to be addressed before decisions about approval or reimbursement, others need to be generated post-approval or after entry into the health system. **Figure 1** below summarises the potential RWE use in each core regulatory review process, from pre- to post-authorisation.

Figure 1. Potential use of real-world evidence in each core regulatory review process

Source: Modified from an original EMA figure.[29]

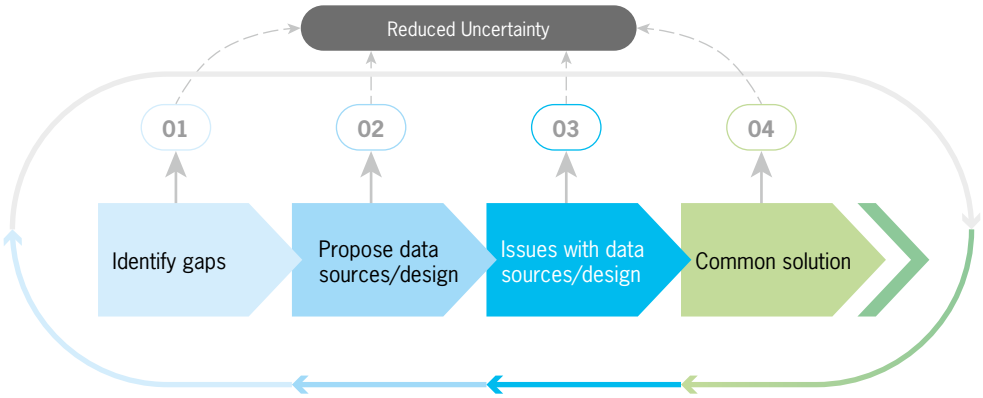


Potential evidence gaps need to be identified by the sponsor early, and agreement on the timing and the type of evidence needed to fill such gaps must be reached early enough to allow sufficient

time to address the research questions.[30,31] It is especially important for the sponsor to deal with gaps in evidence for highly innovative medicines or for rare diseases because of uncertainties about the patient population, the natural history of the disease, the size and durability of clinical effects in comparison to the alternatives, and safety and cost-effectiveness.[32] For example, a framework to identify the gaps in evidence for specialised treatments for rare diseases has been proposed as part of the TRUST4RD tool (**Figure 2**).[33] This framework provides guidance on how to determine the appropriateness and value of filling gaps in evidence with RWD throughout the lifecycle of a medicine as part of a multi-stakeholder collaborative and iterative process. As evidence is generated, uncertainties are reviewed and prioritised, and evidence-generation plans revised or clarified accordingly.

Figure 2. Process proposed by the TRUST4RD Tool

Source:[32] Modified from the original TRUST4RD Tool



Process for matching data sources/designs with evidence gaps

When evidence is generated, the stakeholder needs to review the plan and assess whether or not the evidence generated has answered the research questions, fully or partially, and created new questions to be answered.

The variety of evidence generated, as well as the amount of information derived from it, compel all stakeholders in medicine development to recognise and establish the following:

- Uncertainties may arise and strength of evidence may fluctuate at different decision points (including risk/probability of wrong decision). It is thus important every time new evidence arises, to assess the totality of information and how the new produced information affects the current state of knowledge. The evidence assessment is thus an iterative process as every time new evidence brings new information, ultimately either the evidence gap is narrowed or closed and/or new questions arise.
- The challenges of new evidence emerge throughout a product's lifecycle or after product development. The sponsor must establish a clear and transparent strategy, and an evidence-generation plan must be established, including potential need and frequency of reassessment of the plan every time new information arises. This plan should, ideally, anticipate and adapt to changes in the treatment landscape and new evidence generation. The sponsor's evidence plan should always have the goal of informing the benefit–risk profile of the pharmaceutical product.
- The need for expertise (e.g. RWD/RWE, biostatistics, pharmacoepidemiology) is based on established strategy across all stakeholder groups (pharmaceutical companies, regulators, and payers). Respectful collaboration and open communication among experts across sectors can foster successful outcomes.

1.3.2. Evidence needed to meet stakeholder specific requirements

A strategy for addressing the evidence gaps should cover all types of evidence generation, whether it leads to a clinical trial or an observational study (OS), and should only be based on the research question of interest arising from the evidence needed by different stakeholders. Common stakeholder requirements or expectations are for high-quality data or information and for reliability, access and understanding of the information.

Regulators request at patient population level that the benefits outweigh the risks, taking into account the clinical and regulatory context of the product. To meet regulator's requirements, sponsors provide efficacy/effectiveness and safety data from interventional and/or non-interventional studies combined with information from routine pharmacovigilance activities. Complementarily, RWD can inform on the natural history of the disease, epidemiological features of the disease, unmet medical needs, SOC, and medication utilisation patterns. In addition, RWD allows studying special patient populations, such as paediatric patients, as well as long-term safety and effectiveness.

HTA requires cost effectiveness and budget impact analyses, in addition to the clinical efficacy and safety data. To meet HTA's requirements, sponsors provide cost estimates of the health state, QoL and utilisation of the health state, as well as economic models (e.g. SOC basis computed RCT results). To meet payer's requirements, similar evidence is needed to demonstrate unmet clinical needs, clinical and cost effectiveness, budget impact and health priorities.

To generate evidence that potentially satisfies the needs of both regulators and HTA bodies, EMA offers consultations in parallel with HTA bodies in Europe, allowing medicine developers to obtain feedback from regulators and HTA bodies on their evidence-generation plans to support decision making on marketing authorisation and reimbursement of new medicines at the same time.[34]

Patients and physicians make medical decisions at individual patient level assessing benefits and risks of the treatment of interest. They request evidence on who can benefit the most from the treatment. To meet such requirements, sponsors, regulators, and HTAs/payers provide evidence on diagnostic tools, optimal treatments and SOC, medical history and genetic information of the disease, potential drug-drug interactions, etc.

1.4. Using real-world evidence during product development

RWD and RWE have a key role to play in supporting decision making along the lifecycle of a medicinal product. Whilst they cannot entirely replace the need for controlled experiments such as RCTs, they can complement them at various stages. Strategies that can facilitate and accelerate the medicine development process are of high interest, and regulatory authorities have been evaluating the use of RWD across many stages of the medicine development process.[35] **Figure 3** provides a summary of the various opportunities for RWE generation along the lifecycle.[36,37] In this section, typical applications of RWD in the product lifecycle will be further explored with the help of real-world examples.[38]

Figure 3. Examples of using real-world data in the stages of the medicine lifecycle

Source:[39]

Discovery	Early Development	Full Development	Registration / Market Acess	Lifecycle Management
Identify how many patients currently suffer from the disease alongside comorbidities. Identify current SOC for the disease and assess adherence to clnical guidelines. Identify how many patients with a given disease are insufficiently controlled or have inappropriate treatment for their characteristics.	Identify how current treatment pathways might be affected by a new drug. Identify the costs associated with thue current unmet needs that a new compound might address. Refine inclusion/exclusion criteria of patient cohort to ensure trials better represent the target population.	Use genetic and biomarker information for fast identification of patients meeting inclusion criteria to speed up recruitment to RCTs. Use historical or synthetic controls to replace contemporaneous controls in RCTs, thus reducing cost and number of patients required to participate. Identify rates of adverse events expected during RSCTs.	Assist with health economic evaluation by comparing new compound to current SOC. Identify outcomes associated with current SOC e.g. number of adverse events, disease progression and use of resources, and compare these to new compound. Anticipate budget impact of new compound when introduced across patient segments.	Reduce costs of follow-up trial by using electronic medical records to assess drug effectiveness. Provide information on drug usage within the population e.g. which patients received the drug, dose and duration, patient compliance. Assess outcomes of use e.g. were predicted benefits met? Were the number/type of adverse events as expected?

1.4.1. Real-world evidence during pre-approval phase

RWE use in a product's lifecycle can occur as early as the discovery phase, where researchers investigate the interactions among different molecules, genes, and proteins, with the goal to find novel targets, biomarkers, and compounds.[40] Some of these goals can be achieved using RWD applications. For example, in a review, 20 studies were identified that used RWD to facilitate medicine discovery and clinical research. Among them, 16 identified or validated new phenotypes, disease markers, and biomarkers for patient identification and stratification.[41]

Within early research settings, RWD and RWE can be used to support the discovery of novel targets by identifying unmet medical needs, understanding disease epidemiology and characterising disease burden. They can focus research and development (R&D) efforts by accurately defining the target population, its current standards of care as well as the safety profile of the medications currently used.

During product development, RWD and RWE can be used to design and run clinical trials more efficiently by helping to: (1) identify target patient populations; (2) improve feasibility testing; (3) establish the natural history of disease (particularly for rare diseases); (4) facilitate patient identification and recruitment, clinical site and country selection for global clinical trials; (5) identify disease progression or mortality prognostic biomarkers to inform patient selection for trials (especially oncology medicine development); and (6) accelerate clinical trial execution through novel study designs that make better use of external control arms. Emerging safety issues can be assessed in the light of the natural history of the disease and expected events (background rates) in the population being studied.

Specifically, in the stages of the development phase, RWD can help to:

- Characterise diseases and patient populations, and to understand current unmet medical needs. For example, RWD can estimate how many patients with a given disease have their disease insufficiently controlled or have inadequate treatment and define their characteristics. The RWE can support an orphan drug designation application and paediatric investigation plan development.

- Identify patients for participation in research programmes, which speeds up the recruitment and makes it more efficient. For example, well-managed databases based on EHRs allow fast identification of patients meeting the recruitment criteria of an RCT.
- Represent the population requiring access to the candidate drug, making the design of RCTs more pragmatic (i.e. moving slightly more to the right of the explanatory-pragmatic continuum for trials, to better reflect real life by refining the strict inclusion/exclusion criteria of RCTs). For example, claims databases can show what the numbers of follow-up visits and investigations are in routine practice and this practice can be mimicked in a pragmatic trial.
- Promote inclusion diversity. RWD containing genetic and biomarker information can permit a swifter, more efficient analytical and clinical validation of biomarkers and change the architecture of clinical development programmes (from one protocol for one population with one medicine to multiple combinations). RWD can be obtained through the cross-interrogation of multiple health care records or national biobanks containing genetic and biomarker information, which better enables the identification of target populations and therefore promotes inclusion diversity.
- Sometimes reduce the need for the recruitment of control patients to a randomised trial or single-arm trial through the provision of a synthetic or historical control arm in a time- and cost-efficient manner. For example, RWD collected from sources such as health records, claims data and historical clinical trial data can be used to model a control group that meets the specific requirements of an RCT, thus reducing the need for placebo patients.
- Assess the real-world performance of different diagnostic tests. RWD can be used to facilitate approval for diagnostic testing, such as under emergency use authorisation, as in recent applications during the COVID-19 pandemic.

During the market access phase, RWD can help provide a better understanding of:

- Patient management and modalities of the current SOC for the sake of comparison with the new medicine. For example, in health economic evaluations, the new medicine is typically compared to the SOC. It is therefore vital that the SOC is described as accurately as possible, and differences to be expected in different countries/regions are considered.
- Outcomes in routine clinical practice related to the current SOC, such as the number of complications and adverse reactions, disease progression, resource use and costs.
- Safety issues found during development. RWD can provide the expected background rates of safety events in the target population against which the observed rates of the same events in RCTs can be compared.

Within regulatory submissions and approvals, product developers and regulators are working to understand where and how RWD and RWE can support decision making. RWD and RWE applications are well established for clinical safety and pharmacovigilance monitoring, but more recently have been explored to support new approvals or expanded indications. For example, in the pre-approval phase, RWD from externally controlled trials have been used to support the regulatory approval of new treatments for rare diseases.

During the development phase, RWD can be used to support patient-centred and evidence-driven clinical trials by providing contemporaneous and/or historical control cohorts. We provide detailed case studies in Appendix 1, for example the approval of fosdenopterin (case study A) with external control data from a natural history disease study.

1.4.2. Real-world evidence in post-approval phase

Fulfilling post-approval requirements is generally the area where stakeholders have the most experience using RWD and RWE for regulatory decision making and where regulators have shown

more acceptance. This may be because RWD and RWE are seen as complementary to a large body of evidence already collected during clinical trials.

The following section presents examples of how RWD can be applied to the post-approval phase of the product lifecycle. In the post-approval setting, RWD plays a key role in the benefit–risk assessment of products including (1) long-term and rare adverse safety outcomes; (2) durability of benefit (e.g. duration of vaccine effectiveness or gene therapy); (3) the evaluation of the effectiveness of risk mitigation measures.

After market entry, RWD can help to:

- Provide evidence on the real-world usage of medicines, such as effective dosage or treatment duration and adherence;
- Address safety- or effectiveness-related questions, for example by characterising an identified or potential risk, or by establishing the effectiveness of risk mitigation measures;
- Expand safety-related labelling (e.g. warnings, precautions, or recommended dosing) to enable identification of the need for further monitoring or a specialist visit;
- Support marketing application renewals using comparative effectiveness studies, or benefit–risk balance estimates;
- Support conversion to a full approval, for example by enhancing a confirmatory trial with additional safety and effectiveness data;
- Support a new indication or label extension by using RWD as comparators to the treated patients;
- Characterise special patient sub-populations (e.g. infants or elderly) in terms of safety and efficacy/effectiveness.

For access and reimbursement decisions by HTA/payer, RWD and RWE are used to demonstrate the value of medicinal products and diagnostics for initial access and pricing decisions, support HTA, assess comparative effectiveness, and may provide evidence to support value-based agreements between companies and authorities.

In commercial settings, RWD and RWE are used to monitor and inform customer support programmes and guide commercial strategies, including the competitive landscape and understand patient adherence, switching, and possibly reasons for discontinuation.

In drug utilisation studies, RMPs could employ RWD to evaluate how products are being used to support safe and effective product use, plan for product manufacturing, distribution or discontinuation, and monitor off-label use of medications which may be of value for both medicine safety and medicine repurposing (taking an existing medicine or drug candidate and using it for a medical condition that is different from what it was originally developed to treat).

Patients and HCPs may use RWD and RWE to support treatment decisions. RWE and RWD may be particularly useful in this context when there is an evidence gap, or when questions related to clinical care and QoL may be beyond the scope of clinical trials.

1.5. Evidence generation presentation and communication

For RWE to support regulatory decision making, all stakeholders, including sponsors, regulators, and HTAs need to implement a transparent process of planning, reporting and assessing of RWE.

Transparency of the research processes is key to enable decision makers to evaluate the quality of the methods used and the applicability of the evidence generated. Such transparency will directly improve trust, credibility and reliability in the evidence generated.

Both the ISPE and the ISPOR have actively developed guidance for RWE studies.[42] Best practices include pre-specification of details of the study design and analysis plan and accountability for reproducible research.

1.5.1. Transparency and disclosure of protocol

The structured template for planning and reporting on the implementation of RWE studies (STaRT - RWE) collaborative, a public-private consortium, has developed a structured template for planning and reporting on the implementation of RWE studies of the safety and effectiveness of treatments. The template serves as a guide for: designing and conducting reproducible RWE studies; setting clear expectations for transparent communication of RWE methods; reducing misinterpretation of prose that lacks specificity; allowing reviewers to quickly orient and find key information; and facilitating reproducibility, validity assessment, and evidence synthesis.[43] This information can increase health care decision makers' ability to evaluate RWE studies effectively. HARPER — a protocol template to improve reproducibility — can facilitate study protocol development and enhance transparency and reporting.[44]

In addition, to enhance transparency in RWD research, numerous public repositories exist for the registration of RWE protocols for future inspection, including the European post-authorisation study (EU PAS) Register,[45] clinicaltrials.gov, and Health Services Research Projects (HSRProj). The EU PAS Register has a source data repository that provides information also on the source of the data.[46]

While transparency and disclosure are needed for evaluation, it is also the researchers' responsibility to communicate study results unambiguously, including providing a critical assessment of the evidence produced. In that respect, leveraging existing methodology (ICH M4E) to present RWE to regulators using the full extent of clinical overview and the effects tables from structured benefit–risk assessment, summarising the existing evidence, and re-stating the rationale for the new study (with context), highlighting uncertainties and limitations of the research methods, also explicitly contextualises results.[47] The inclusion of assessment of RWD in an effects table can make it explicit what value is added, and it can serve to build trust on reported RWE and establish the need for further investigations.

1.5.2. Cross-stakeholder interaction and collaboration

The need for discussion and consensus by multiple stakeholders for the acceptability of plans to generate RWD/RWE has recently been highlighted.[48] For example, the European Network for Health Technology Assessment (EUnetHTA) started in 2006 as an informal initiative, supporting collaboration between European HTA organisations, bringing value at the European, national, and regional levels through the facilitation of efficient HTA resource use, the creation of a sustainable system of HTA knowledge sharing, and the promotion of good practice in HTA methods and processes.[49]

Since 2017, EMA and EUnetHTA have offered parallel advice services called Early Dialogues to provide a platform for multi-stakeholder interactions.[50] This parallel consultation by regulators and HTAs offered sponsors opportunities for mutual understanding and problem solving between regulators and HTAs, the goal being to facilitate robust evidence generation for different stakeholders.[51]

Another example of collaboration between regulators and HTAs has been the formal recognition by the EMA of the EUnetHTA Registry Evaluation and Quality Standards Tool (REQuest).[52] More recently, EUnetHTA has evolved into the Regulation on Health Technology Assessment (HTAR), which entered into force in 2022 as the permanent system of HTA collaboration in Europe.[53] A parallel submission process by Health Canada, CADTH and Quebec's Institut national d'excellence en santé et en services sociaux (INESSS) was established in 2018.

While such initiatives have not been specifically created with RWE/RWD in mind, they provide an early opportunity for different stakeholders to discuss the appropriateness and acceptability of RWE. CADTH and NICE also offer sponsors joint scientific advice upon request.[54]

1.6. Engaging with regulators

Most regulatory agencies encourage early discussion through transparent information sharing and meeting requests.

1.6.1. US Food and Drug Administration

In response to the 21st Century Cures Act, the US FDA has developed a RWE Program and provided guidance on RWE use for regulatory decision making. RWE can be submitted to the US FDA in an investigational new drug (IND), an biologics license application (BLA) or new drug application (NDA) submission or a meeting request, with a cover letter indicating that the submission contains RWE.[55] RWE may be submitted in at various phases of the lifecycles of product development. For example, RWE may be submitted in an IND phase to examine the natural history of disease using RWD, or in a NDA/BLA application to provide external controls for a single arm trial, or in a post-marketing phase to fulfil a post-approval requirement to further evaluate safety or effectiveness.

Early communications between the US FDA and sponsors are critical for RWE use for regulatory purposes.

1.6.2. European Union and European Economic Area

The European medicines regulatory system is based on a network of around 50 regulatory authorities from the 30 European Economic Area countries (27 EU Member States plus Iceland, Liechtenstein and Norway), the European Commission and EMA.

EU regulators use RWD analysis in the post-approval phase on a regular basis, mostly to further characterise safety, but also that for effectiveness.[56] During the pre-approval phase, the evidence generated from RWD has been seen to complement the evidence from RCTs.[57] There is however increasing interest in the use of RWD to support regulatory decision making across the product lifecycle.[58,59]

Scientific advice[60] is given by the Committee for Medicinal Products for Human Use (CHMP) on the recommendation of the Scientific Advice Working Party (SAWP). Of note, the EMA has a programme to provide parallel scientific advice (PSA) to sponsors.

The EMA also offers consultations in parallel with the HTA bodies in Europe as of 2017. This aims to allow medicine developers to obtain feedback from regulators and HTA bodies on their evidence-generation plans to support decision making on marketing authorisation and reimbursement of new medicines at the same time. This initiative is also of value for testing the fitness of RWD- and RWE-related proposals to address the expectations of different public stakeholders.

The conditions of successful pre- or peri-approval use of RWE in the EU regulatory approval process have thus far been related to: the rarity of disease/orphan indication; context of significant unmet need; important need for rapid access to medicines; the impracticality of performing a RCT; or other challenges of following the traditional medicine development pathway.[61]

In 2023, EMA published a real-world evidence framework to support EU regulatory decision-making, reporting on the experience gained with regulator-led studies since 2021.[62]

Furthermore, in April 2024, EMA published a draft reflection paper on the use of RWD in non-interventional studies to generate RWE.[63]

1.6.3. Other national regulators

Australia

Following a review into the usage of RWE and PROs, the Therapeutic Goods Administration (TGA) has made changes to its regulation of medicines and medical devices.[64]

TGA has adopted a definition of RWE and revised its pre-submission planning form for medicines to include RWE usage declaration in the application. At the same time the clinical evidence guidelines for medical devices have been updated. TGA has also adopted international scientific guidelines on RWE such as those from EMA and US FDA.

TGA is clear that the potential for RWE is increasing as various forms of health data accumulate and new analytical frameworks are developed.

Brazil

The Brazilian Health Regulatory Agency (Anvisa) has been seeking to increase knowledge on RWD and RWE use for regulatory decision making. The Agency has promoted technical discussions with several stakeholders, such as academic institutions, pharmaceutical companies, and regulators. At these meetings, discussions covered potential options for:

- The collection, quality control, validation, and acceptability of RWD;
- Information on initiatives from other regulatory agencies on this topic;
- Case studies of pharmaceutical companies and use of RWE at different stages of clinical medicine development;
- Data analysis driven by artificial intelligence in health care settings;
- Opportunities and challenges of RWE studies; and
- Perspectives of medical professionals and industry in relation to RWE.[65]

Anvisa has begun its internal process of building understanding for the critical assessment of RWD and RWE. Several key aspects should be discussed with Anvisa prior to submission if there is an intent to use RWD and RWE to support claims of efficacy and safety, especially for medicines aimed at treating rare diseases and serious and debilitating conditions. Such discussions should cover, for example: pertinence of using primary or secondary sources of RWD; use of national or international data sources; uncertainties related to outcomes, follow up, sample size, comparators, and target population; and design of studies that include RWD.[66]

This communication can be established through the following existing channels: pre-submission meetings for scientific advice (available for the medicine registration process, post-approval changes,

and clinical research for regulatory purposes); discussions of queries issued by Anvisa (for ongoing reviews); and ombudsman systems (which can be used not only by the pharmaceutical companies, but also by citizens and other government departments interested in seeking clarity from the Agency).[67]

Current strategies will contribute to the improvement of the current model of generating information, focusing on the participant/patient. The initiative called Digital Health Strategy, which will include the National Health Data Network (Rede Nacional de Dados em Saúde (RNDS), a component of the national health database, will seek integration and interoperability of health information not only between public and private health institutions, but also among health management departments of federal entities to ensure access to health information required for the continuity of participant/patient care. RNDS information may be valuable for epidemiological, statistical, research, and regulatory purposes.

To encourage the interoperability of health data through publication in a machine-processable format and to promote the continuous improvement of the quality of the data made available, Anvisa also developed an inventory of the databases under its custody for public awareness of databases maintained by the Agency. This initiative is called Anvisa's open data plan.[68] With the publication of the Anvisa's open data and the availability of qualified data to society, Anvisa takes an important step towards transparency and social control (i.e. rules and standards in society that keep individuals bound to conventional standards), in line with the principles of publicity and efficiency for regulatory decision making.[69]

Canada

Health Canada, the Canadian government's public health department, continues to optimise the use of RWE for regulatory decisions in order to improve the extent and rate of access to prescription medicines in Canada. This work is being conducted in partnership with key organisations, including the CADTH and the Institut national d'excellence en santé et en services sociaux (INESSS), as described in Health Canada's Notice on Optimizing the Use of Real World Evidence to Inform Regulatory Decision-Making. Health Canada also posted an accompanying document on principles to consider on the generation of high quality RWD in Elements of Real World Data/Evidence Quality throughout the Prescription Drug Product Life Cycle.

Health Canada continues to collaborate with domestic and international partners on the use of RWD/RWE in regulatory decision making and to develop and harmonise RWD/RWE reporting standards. For example, CADTH, Health Canada and other stakeholders developed the Guidance for Reporting RWE to Support Decision-making. More recently, Health Canada has reiterated its position on RWD and RWE by issuing a statement on Health Canada's position on the CADTH Guidance for Reporting RWE to Support Decision-making.

Japan

The Pharmaceuticals and Medical Devices Agency (PMDA) uses RWD/RWE mainly for safety assessment in the post-approval setting.[70]

PMDA launched the Medical Information for Risk Assessment Initiative (MIHARI) project in 2009 with the aim of strengthening post-approval safety measures for pharmaceuticals.[71] In the MIHARI Project, PMDA has conducted safety assessments of medicines using pharmacoepidemiological methods, with secondary use of electronic medical information that hospitals enter and accumulate for routine medical care, such as claims data and electronic medical records (EMRs).

For example, many pharmacoepidemiological studies have been conducted based on RWD from the National Claims Database (NDB) in Japan[72,73,74] and Medical Information Database NETwork (MID-NET®)[75,76,77] a reliable and valuable database operated and managed by the PMDA in Japan.[78]

Some of those results have led to safety measures such as a revision of precautions on the package insert in Japan.[79,80]

To further improve post-approval pharmacovigilance in Japan, the GPSP (Good Post Marketing Study Practice), which sets reliability standards for post-approval study conducted by MAHs after medicine approval, were revised in 2017.[81] With this revision, post-approval database study has been clearly defined in Japan for promoting RWD utilisation for regulatory purpose.

Japan Revitalization Strategy revised in 2016 (Cabinet decision on June 2, 2016) announced the decision to promote development in Japan by constructing novel clinical development methodologies, more specifically, constructing the disease registry system and thereby proceeding with construction of the clinical innovation network (CIN) that develops clinical development infrastructure based on the disease registry information.[82] Since then, joint industry-academia R&D projects that use the registries have been supported.

The registry utilisation for evaluating safety and efficacy of medicines and medical devices was clarified in the conditional accelerated approval system for medicines and medical devices, which has been started in 2017.[83,84] In 2021, the Ministry of Health, Labour and Welfare (MHLW) published two notifications to promote regulatory use of the registry as follows: basic principles on registry utilisation,[85] and point to consider for assurance of the reliability of using registry data as approval applications.[86]

In addition, the PMDA started activities on Projects Across Multi-Offices, and set up a RWD Working Group in April 2021 to discuss all participants on RWD comprehensively including general principles on RWD utilisation and data reliability in regulatory settings.[87]

See case study E on Cardiovascular risk of urate-lowering drugs: a study using the National Database of Health Insurance Claims and Specific Health Check-ups of Japan and case study F on Nested case-control study utilising MID-NET® on thrombocytopenia associated with pegfilgrastim in patients treated with antineoplastic agents.

People's Republic of China

The importance of RWE in clinical and regulatory decision making has been increasingly recognised in China, with policies and guidelines published in recent years. In January 2020, the National Medical Products Administration (NMPA) of the People's Republic of China published *Guidance on Real-World Evidence Supporting Drug Development and Review (Pilot)*, which outlined the definition and sources of RWD and provided guidance on using RWE in supporting medicine review, indication expansion, post-approval evaluation, and R&D of traditional Chinese medicine. Following publication of that guidance, a technical guideline on the development and review of medicines for children was released in September 2020 by the Centre for Drug Evaluation, an affiliated institution of the NMPA. Besides medicines, RWD are also used in the clinical evaluation of medical devices, for which a technical guideline was published by the NMPA in November 2020.

Chapter 1 – References

1 Definition from techopedia. What is interoperability? (Website accessed 28 October 2022).
2 Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU—actions for stakeholders. *International Journal of Technology Assessment in Health Care*. 2020;36(4):459-468. https://doi.org/10.1017/S026646232000063X.
3 International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Final Concept Paper. Establishment of a new ICH guideline on "General principles on plan, design, and analysis of pharmacoepidemiological studies that utilize real-world data for safety assessment of medicines", 23 March 2022, Endorsed by the Management Committee on 5 April 2022. (PDF accessed 26 February 2024).
4 O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: a milestone in international collaboration. *International Journal of Technology Assessment in Health Care*. 2020;36(3):187-190. https://doi.org/10.1017/S0266462320000215.
5 O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*. 2020;36(3):187-190. https://doi.org/10.1017/S0266462320000215.
6 Bartlett SJ, Barnes T, McIvor RA. Integrating patients into meaningful real-world research. *Annals of the American Thoracic Society*. 2014;11(Supplement 2):112-117. https://doi.org/10.1513/AnnalsATS.201309-327RM.
7 The United States Food and Drug Administration (FDA). Framework for FDA's Real-World Evidence Program. (PDF accessed 15 October 2022).
8 Patient Involvement in the development, regulation and safe use of medicines. CIOMS Working Group XI report. Geneva: Switzerland: Council of International Organizations of Medical Sciences (CIOMS), 2022. https://doi.org/10.56759/iiew8982 (PDF accessed 20 April 2023).
9 Patient involvement in the development, regulation and safe use of medicines. CIOMS Working Group report. Geneva, Switzerland: Council for International Organizations of Medical Sciences (CIOMS), 2022. https://cioms.ch/publications/product/patient-involvement/. Modified from: European Medicines Agency, About us, Glossary of regulatory terms: 'Marketing authorisation holder'. (Webpage accessed 21 March 2024).
10 International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Reflection Paper. International Harmonisation of Real-World Evidence Terminology and Convergence of General Principles Regarding Planning and Reporting of Studies Using Real-World Data, with a Focus on Effectiveness of Medicines Under public consultation until 30 September 2023, Endorsed by the ICH Assembly on 13 June 2023. (PDF accessed 26 February 2024).
11 Purpura, C.A., Garry, E.M., Honig, N., Case, A. & Rassen, J.A. The role of real-world evidence in FDA-approved new drug and biologics license applications. *Clin. Pharmacol. Ther*. 111, 135–144 (2022). https://ascpt.onlinelibrary.wiley.com/doi/10.1002/cpt.2474.
12 The United States Food and Drug Administration (FDA). Real-World Data and Real-World Evidence are Playing an Increasing Role in Health Care Decisions. (Website accessed 15 October 2022).
13 The United States Food and Drug Administration (FDA). Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. Guidance for Industry. (PDF accessed 17 April 2023).
14 The United States Food and Drug Administration (FDA). Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products. Guidance for Industry. (PDF accessed 17 April 2023).
15 The United States Food and Drug Administration (FDA). Data Standards for Drug and Biological Product Submissions Containing Real-World Data. Guidance for Industry. (PDF accessed 17 April 2023).
16 The United States Food and Drug Administration (FDA). Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. Draft Guidance for Industry. (PDF accessed 17 April 2023).
17 The United States Food and Drug Administration (FDA). Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products. Guidance for Industry. (PDF accessed 17 April 2023).
18 The United States Food and Drug Adminstration (FDA). Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. (PDF accessed 2 May 2023).
19 European Medicines Agency (EMA). Scientific Guidance on Post-Authorisation Efficacy Studies. (Website accessed 15 October 2022).
20 European Medicines Agency (EMA). Guideline on Registry-Based Studies. (Website accessed 15 October 2022).
21 Heads of Medicines Agencies (HMA). HMA-EMA Joint Big Data Taskforce Phase II Report: 'Evolving Data-Driven Regulation'. (PDF accessed 15 October 2022).
22 Heads of Medicines Agencies (HMA). Priority Recommendations of the HMA-EMA Joint Big Data Task Force. (PDF accessed 15 October 2022).
23 Preparing to Engage with Health Authorities on Real World Evidence for Medicines: A TransCelerate Perspective Copyright © 2022 TransCelerate BioPharma Inc. All Rights Reserved. (PDF accessed 4 June 2023).
24 Makady A, van Veelen A, Jonsson P, Moseley O, D'Andon A, de Boer A, et al. Using real-world data in Health Technology Assessment (HTA) practice: a comparative study of five HTA agencies. *Pharmacoeconomics*. 2018;36(3):359-368. https://doi: 10.1007/s40273-017-0596-z.
25 International Society for Pharmacoepidemiology (ISPE). Strategic Initiatives: RWE for Regulatory Decision Making. (Website accessed 15 October 2022).
26 Canada's Drug and Health Technology Agency (CADTH). Use of Real-World Evidence in Single-Drug Assessments Environmental Scan. (Website accessed 15 October 2022).
27 NICE real-world evidence framework. (Website accessed 23 May 2023).
28 Bullement A, Podkonjak T, Robinson MJ, Benson E, Selby R, Hatswell AJ, et al. (2020). Real-world evidence use in assessments of cancer drugs by NICE. *International Journal of Technology Assessment in Health Care*. 2020;10;1-7. https://doi.org/10.1017/S0266462320000434.

29 DARWIN EU (Data Analytics and Real World Interrogation Network) PCWP and HCPWP Data workshop 23 September 2020. Peter Arlett EMA Head of EMA Data Analytics and Methods Taskforce Co-chair HMA-EMA Big Data Steering Group. Slides 7-9. (PDF accessed 2 June 2023).
30 Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU—actions for stakeholders. *International Journal of Technology Assessment in Health Care*. 2020;36(4):459-468. https://doi.org/10.1017/S026646232000063X.
31 Annemans L, Makady A. TRUST4RD: tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. *Orphanet Journal of Rare Diseases*. 2020;15(1):1-9. https://doi.org/10.1186/s13023-020-01370-3.
32 Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU—actions for stakeholders. *International Journal of Technology Assessment in Health Care*. 2020;36(4):459-468. https://doi.org/10.1017/S026646232000063X.
33 Annemans L, Makady A. TRUST4RD: tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. *Orphanet Journal of Rare Diseases*. 2020;15(1):1-9. https://doi.org/10.1186/s13023-020-01370-3.
34 EMA Parallel joint scientific consultation with regulators and health technology assessment bodies. (Website accessed 2 June 2023).
35 Chen Z, Liu X, Hogan W, Shenkman E, Bian J. Applications of artificial intelligence in drug development using real-world data. *Drug Discovery Today*. 2021;1;26(5):1256-1264. https://doi.org/10.1016/j.drudis.2020.12.013 (Journal abstract accessed October 2022).
36 Orpha Strategy. Patient-Centric Real-World Evidence. (Website accessed October 2022).
37 Bate A, Juniper J, Lawton AM, Thwaites RM. Designing and incorporating a real world data approach to international drug development and use: what the UK offers. international drug development and use: what the UK offers. *Drug Discovery Today*. 2016;1;21(3):400-405. https://doi.org/10.1016/j.drudis.2015.12.002.
38 Institut national d'assurance maladie-invalidité (INAMI). Annemans L. The use of real world data throughout an innovative medicine's lifecycle. (Website accessed 15 October 2022).
39 Bate A, Juniper J, Lawton AM, Thwaites RM. Designing and incorporating a real world data approach to international drug development and use: what the UK offers. international drug development and use: what the UK offers. *Drug Discovery Today*. 2016;1;21(3):400-405. https://doi.org/10.1016/j.drudis.2015.12.002.
40 The United States Food and Drug Administration (FDA). The Drug Development Process. (Website accessed October 2022).
41 Singh G, Schulthess D, Hughes N, Vannieuwenhuyse B, Kalra D. Real world big data for clinical research and drug development. *Drug Discovery Today*. 2018;1;23(3):652-660. https://doi.org/10.1016/j.drudis.2017.12.002.
42 International Society for Pharmacoepidemiology (ISPE). Resources: Guidelines for Good Pharmacoepidemiology Practices (GPP). (Website accessed 15 October 2022).
43 Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ*. 2021;12;372. https://doi.org/10.1136/bmj.m4856.
44 Wang SV, Pottegård A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, et al. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force. *Value in Health*. 2022;1;25(10):1663-1672. https://doi.org/10.1016/j.jval.2022.09.001.
45 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP). The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). (Website accessed 15 October 2022).
46 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP). ENCePP Resources Database. (Website accessed 15 October 2022).
47 The United States Food and Drug Administration (FDA). Considerations for the Use of Real-World Data and Real- World Evidence to Support Regulatory Decision-Making for Drug and Biological Products. Guidance for Industry. Draft Guidance. (PDF accessed 8 April 2023).
48 Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU—actions for stakeholders. *International Journal of Technology Assessment in Health Care*. 2020;36(4):459-468. https://doi.org/10.1017/S026646232000063X.
49 European Network for Health Technology Assessment (EUnet HTA). About EUnetHTA. (Website accessed 15 April 2023).
50 European Medicines Agency (EMA). Parallel Consultation with Regulators and Health Technology Assessment Bodies. (Website accessed 15 October 2022).
51 European Network for Health Technology Assessment (EUnetHTA). Guidance on parallel EMA/EUnetHTA 21 Joint Scientific Consultation. (PDF accessed 15 October 2022).
52 *European Medicines Agency (EMA). Guideline on Registry-Based Studies*. (PDF accessed 15 October 2022).
53 European Network for Health Technology Assessment (EUnetHTA). Guidance on parallel EMA/EUnetHTA 21 Joint Scientific Consultation. (PDF accessed 15 October 2022). Regulation on Health Technology Assessment (Website accessed 26 February 2024).
54 Government of Canada. Notice to Industry: Aligned Reviews between Health Canada and Health Technology Assessment Organizations. (Website accessed 15 October 2022).
55 The United States Food and Drug Administration (FDA). Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drugs and Biological Products. Guidance Document. (PDF accessed 15 October 2022).
56 European Medicines Agency (EMA). Scientific Guidance on Post-Authorisation Efficacy. (PDF accessed 15 October 2022).
57 European Medicines Agency (EMA). Regulatory Perspective on Real-World Evidence (RWE) in Scientific Advice. (PDF accessed 15 October 2022).
58 Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clinical Pharmacology and Therapeutics*. 2019;106(1):36-39. https://doi.org/10.1002/cpt.1426.
59 Eichler HG, Pignatti F, Schwarzer-Daum B, Hidalgo-Simon A, Eichler I, Arlett P, et al. Randomized controlled trials versus real world evidence: neither magic nor myth. *Clinical Pharmacology & Therapeutics*. 2021;109(5):1212-1218. https://doi.org/10.1002/cpt.2083.
60 European Medicines Agency (EMA). Human Regulatory: Scientific Advice and Protocol Assistance. (Website accessed 15 October 2022).

61 Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clinical Pharmacology and Therapeutics*. 2019;106(1):36-39. https://doi.org/10.1002/cpt.1426.
62 European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making, Report on the experience gained with regulator-led studies from September 2021 to February 2023. EMA/289699/2023. (PDF accessed 7 March 2024).
63 European Medicines Agency. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence. Draft. 15 April 2024. EMA/CHMP/150527/2024. (PDF accessed 7 May 2024).
64 Australian Government Department of Health and Aged Care Therapeutic Goods Administration. Real World Evidence (RWE) and patient reported outcomes (PROs). (Website accessed 12 April 2024).
65 Brazilian Health Regulatory Agency (Anvisa). Falcão MG, Costa PL, Soares RL, Santos GML. RWD and RWE in Brazil: Perspective of the Brazilian Health Regulatory Agency (Anvisa). DIA Global Forum. 2020;11. (Website accessed 2 May 2023).
66 Brazilian Health Regulatory Agency (Anvisa). Falcão MG, Costa PL, Soares RL, Santos GML. RWD and RWE in Brazil: Perspective of the Brazilian Health Regulatory Agency (Anvisa). DIA Global Forum. 2020;11. (Website accessed 2 May 2023).
67 Brazilian Health Regulatory Agency (Anvisa). Falcão MG, Costa PL, Soares RL, Santos GML. RWD and RWE in Brazil: Perspective of the Brazilian Health Regulatory Agency (Anvisa). DIA Global Forum. 2020;11. (Website accessed 2 May 2023).
68 Carvalho-Soares ML, Hemesath-Mendes KS, Ribeiro AM, Andrade AC, Vidal MB. Open data for society: a history of the National Health Surveillance Agency - Anvisa Plan. *Brazilian Journal of Hospital Pharmacy and Health Services*. 2021;12(2):1-4. https://doi.org/10.30968/rbfhss.2021.122.0647.
69 Carvalho-Soares ML, Hemesath-Mendes KS, Ribeiro AM, Andrade AC, Vidal MB. Open data for society: a history of the National Health Surveillance Agency - Anvisa Plan. *Brazilian Journal of Hospital Pharmacy and Health Services*. 2021;12(2):1-4. https://doi.org/10.30968/rbfhss.2021.122.0647.
70 Nishioka K, Makimura T, Ishiguro A, Nonaka T, Yamaguchi M, Uyama Y. Evolving acceptance and use of RWE for regulatory decision making on the benefit/risk assessment of a drug in Japan. *Clinical Pharmacology & Therapeutics*. 2022;111(1):35-43. https://doi.org/10.1002/cpt.2410.
71 Ishiguro C, Takeuchi Y, Uyama Y, Tawaragi T. The MIHARI project: establishing a new framework for pharmacoepidemiological drug safety assessments by the Pharmaceuticals and Medical Devices Agency of Japan. *Pharmacoepidemiology and Drug Safety*. 2016;25(7):854-859. https://doi.org/10.1002/pds.4032.
72 Komamine M, Kajiyama K, Ishiguro C, Uyama Y. Cardiovascular risks associated with dipeptidyl peptidase-4 inhibitors monotherapy compared with other antidiabetes drugs in the Japanese population: a nationwide cohort study. *Pharmacoepidemiology and Drug Safety*. 2019;28(9):1166-1174. https://doi.org/10.1002/pds.4847.
73 Shida H, Kajiyama K, Sawada S, Ishiguro C, Kubo M, Kimura R, et al. Use of national database of health insurance claims and specific health checkups for examining practical utilization and safety signal of a drug to support regulatory assessment on postmarketing drug safety in Japan. *Frontiers in Medicine*. 2023;10. https://doi.org/10.3389%2Ffmed.2023.1096992.
74 Sawada S, Kajiyama K, Shida H, Kimura R, Nakazato Y, Iguchi T. Cardiovascular risk of urate-lowering drugs: a study using the National Database of Health Insurance Claims and Specific Health Checkups of Japan. *Clinical and Translational Science*. 2023;16(2):206-215. https://doi.org/10.1111/cts.13439.
75 Yamada K, Itoh M, Fujimura Y, Kimura M, Murata K, Nakashima N, et al. The utilization and challenges of Japan's MID-NET® medical information database network in postmarketing drug safety assessments: A summary of pilot pharmacoepidemiological studies. *Pharmacoepidemiology and Drug Safety*. 2019;28(5):601-608. https://doi.org/10.1002/pds.4777.
76 Hasegawa T, Sawada S, Ishiguro C, Ando T, Kobayashi K, Komiyama N, et al. Assessing the risk of decrease in kidney function in patients prescribed direct-acting antivirals for hepatitis C utilizing the MID-NET® Medical Information Database Network in Japan. *Therapeutic Innovation & Regulatory Science*. 2022;56(4):625-631. https://doi.org/10.1007/s43441-022-00400-5.
77 Kinoshita Y, Kajiyama K, Ishiguro C, Nonaka T, Kimura R, Kikuchi Y, et al. Characterizing granulocytopenia associated with thiamazole in patients with hyperthyroidism based on real-world data from the MID-NET in Japan. *Clinical Pharmacology & Therapeutics*.2023;113(4):924-931. https://doi.org/10.1002/cpt.2850.
78 Yamaguchi M, Inomata S, Harada S, Matsuzaki Y, Kawaguchi M, Ujibe M, et al. Establishment of the MID-NET® medical information database network as a reliable and valuable database for drug safety assessments in Japan. *Pharmacoepidemiology and Drug Safety*. 2019;28(10):1395-1404. https://doi.org/10.1002/pds.4879.
79 Kajiyama K, Ishiguro C, Ando T, Kubota Y, Kinoshita N, Oniyama Y, et al. Nested case-control study utilizing MID-NET® on thrombocytopenia associated with pegfilgrastim in patients treated with antineoplastic agents. *Clinical Pharmacology & Therapeutics*. 2021;110(2):473-479. https://doi.org/10.1002/cpt.2263.
80 Sawada S, Ando T, Hirano M, Komiyama N, Iguchi T, Oniyama Y, et al. Effect of hepatitis C drugs on blood coagulability in patients on warfarin using the Medical Information Database Network (MID-NET®) in Japan. *Therapeutic Innovation & Regulatory Science*. 2021;55:539-544. https://doi.org/10.1007/s43441-020-00247-8.
81 Ministry of Health, Labour and Welfare. Amendment of ministrial ordinance (no 116). Good Postmarketing Study Practice (original ordinance No 171, 2004). (PDF accessed on 2 May 2023).
82 Japan Revitalization Strategy 2016. (PDF accessed on 2 May 2023).
83 Ministry of Health, Labour and Welfare, Pharmaceutical Safety and Environmental Health Bureau. Conditional early approval system for innovative medical device products (Fast-Break Scheme) (PSEHB Notification No. 0731-1, July 31, 2017). (PDF accessed 2 May 2023).
84 Ministry of Health, Labour and Welfare, Pharmaceutical Safety and Environmental Health Bureau. Implementation of a conditional early approval system for pharmaceutical products (PSEHB/PED Notification No. 1020-1, October 20, 2017). (PDF accessed 2 May 2023).
85 Ministry of Health, Labour and Welfare, Pharmaceutical Safety and Environmental Health Bureau, Basic principles on utilization of registry for Applications (PSEHB/PED Notification No.0323-1, PSEHB/MDED Notification No.0323-1, March 21, 2021). (PDF accessed 2 May 2023).
86 Ministry of Health, Labour and Welfare, Pharmaceutical Safety and Environmental Health Bureau. Points to consider for ensuring the reliability in utilization of registry data for applications (PSEHB/PED Notification No.0323-2, PSEHB/MDED Notification No.0323-2, March 23, 2021). (PDF accessed 2 May 2023).
87 Pharmaceuticals and Medical Devices Agency (PMDA). Projects Across Multi-Offices in PMDA. RWD WG. (Website accessed 2 May 2023).

CHAPTER 2.
SOURCES OF REAL-WORLD DATA

The scope of RWD usually includes health care data sources that can provide information that can be used to infer the benefits and risks of medicinal products and measure resource utilisation. While this scope is appropriate, it is incomplete. There are other sources that, although not as rich for capturing information from the provision of health care, are useful for evaluating the safety and effectiveness of products and the burden of diseases in different populations. They include spontaneous reporting systems (SRSs) and surveys. Such sources have been used to evaluate the benefits and risks of products for decades. For the purpose of this document, health care data and other sources mentioned above will be called traditional data sources.

The introduction of new technologies such as those related to remote care and the increased use of mobile devices has provided new sources of information that can be generated with unprecedented volume, speed and complexity, and require a different set of data management and analytical methods. Although current use of these emerging sources is still limited compared to the traditional ones, with the rapid development of modern computing and advanced analytics, it is just a matter of time before they will also be used as key RWD sources in the context of regulatory decision making.

This chapter describes both traditional and emerging data sources, focusing on key features for the purpose of various regulatory uses.

2.1. Traditional sources

2.1.1. Existing health care databases

RWD from existing health care databases, including insurance claims, EHR, and registry databases have been used for decades — often referred to as secondary use/analysis, mainly for safety evaluation, risk management and to support benefit–risk evaluation of medicinal products.[1,2] These uses are widely accepted, and these data have many strengths. They are longitudinal in nature, with records of the same patients being available at different points in time, and thereby enable the establishment of a temporal sequence, which is essential in the evaluation of a causal relationship. Other strengths include:

- The population size and the number at risk (patients without events at baseline) can be clearly defined. This combined with the ascertainment of the number of events during the follow-up allow estimations of risks. The size of the exposed population at risk is not available in SRSs, and therefore risks cannot be estimated.
- Comparison groups, or comparators, are often available to evaluate potential associations between medicinal products' benefits and risks, and can be more easily defined than in other types of databases. In addition, different types of comparison groups can be assembled more easily than in clinical trials.
- Although perhaps not complete, these databases usually contain information on a large number of potential confounders such as demographic characteristics, comorbidities, and concomitant medications.

- Many of these databases allow analyses of much larger numbers of individuals than available in clinical trials, and with longer follow-up. Thus, they are more suitable for identifying rare events that may not be detected in smaller clinical trials.
- The data are captured from a real-life health care setting, making results more generalisable to the target population, and offering an opportunity for analyses in different sub-groups not available in trials, such as older adults, pregnant women, children, and different racial and ethnic groups, as well as permitting examination of off-label use.
- Because the data are already available, these studies can be conducted in less time than analogous studies requiring ad-hoc data collection.

Existing health care databases also have limitations. Some of them, especially insurance claims and medical records databases, are created for reasons other than research and, consequently, may not be suitable for answering certain questions. For example, many claims databases in the US have incomplete or no information on death, precluding their use to study mortality as an endpoint, although improvement has been made by linking them to death registries.

There are several issues to consider when directly linking data including confidentiality and ethical issues. Other limitations include:

- The time lag between occurrence of the events and the availability of the health care databases for analyses, which is problematic especially when evaluating the effectiveness of risk minimisation measurements in real time.
- For some sources, populations are limited to patients with certain conditions or exposed to specific medicinal products. Consequently, they are not suitable for other conditions or medicinal products. This is more of a problem for specific registries than claims or EHR databases where patients with different diseases and/or exposed to different medicinal products are included.
- The availability of information on potentially important confounding variables and how availability differs across different health care databases. For example, body mass index (BMI), smoking, and laboratory values may be available in EHRs but missing from many insurance claims databases.
- The availability of information for certain sub-groups of important patient populations (for example, older adults, children, and pregnant women) may also vary.
- The validity of information both of exposures and outcomes may not be ideal. Validity could be evaluated and improved by comparing information in the RWD sources to that in other sources that could be considered gold standards.[3,4]

Given these limitations, when performing new studies, it is important to involve or consult the parties who are close to the development of those health care databases, or who have had experience in using them, ideally during the whole study period. This will ensure the appropriate use of the data elements (including coding system and outcomes definitions), study designs and methods, to answer the study questions.

In addition to the limitations related to the characteristics of the health care databases mentioned above, there are also other challenges related to the approach to analysing them. One methodological issue that has been discussed for a long time is repeated analyses. A health care database may be used for multiple analyses of the same outcome by different parties or at different points in time. It may also be used for analyses of many different outcomes. If the hypothesis testing in the RWD study is to serve as basis for a regulatory decision, should multiple testing be addressed? Some suggest adjustment is not necessary because it is not a clinical trial[5] while others prefer some kind of adjustment.[6]

To date, health care databases have been mainly used to address safety issues such as the evaluation of a finite number of hypotheses that have been set a priori (hypothesis testing or signal evaluation)

and confirmation of potential safety issues identified in other data sources (signal confirmation or refinement)[7,8] but less commonly for signal detection with no a priori hypothesis.[9] Besides the challenges already mentioned above, the use of the same database both for signal detection and signal evaluation presents another challenge.[10] Some suggest that signal detection (or hypothesis generation) should be done independently from signal evaluation (or hypothesis testing) in a different data source.[11] Others suggest that the two could be done in the same databases as long as the methods of analyses are different.[12,13]

Although the use of health care databases for RWD studies on benefits (effectiveness) has been limited and more controversial[14] there has been a lot of discussion on how they can be used as part of regulatory decision making. Many reasons for scepticism by regulators have already been discussed above. To date, RCTs are still considered the gold standard for assessment of benefit (efficacy and effectiveness) and, other factors being equal, of being less prone to many of the biases to which OSs are prone, especially for new products.[15]

All case studies A – H in Appendix 1 used existing RWD sources, with case study A using a combination of trial data, an existing RWD source, and RWD with primary data collection.

2.1.2. Real-world data sources with primary data collection

When existing data sources are not suitable to answer the questions at hand, either due to lack of information or differences in the study populations, a new RWD with primary data collection, or with data collected on an ad-hoc basis specifically for the study, is needed. For example, some existing data sources lack information connecting mothers and their babies and, therefore, are not suitable for evaluating association between exposure to medicinal products used during pregnancy and the pregnancy outcomes. For this purpose, a pregnancy registry with primary data collection can be done and designed specifically to test a particular set of hypotheses about the pregnancy outcomes of interest.

Another example is when the US FDA issued guidance on cardiovascular safety issues possibly related to new diabetes medicines among patients with type 2 diabetes.[16] This guidance led to new studies, including non-randomised/RWD studies that extended the evidence from pre-approval clinical trials to post-marketing real-world settings.

RWD sources with primary data collections have their advantages and limitations. Ad-hoc data collection is performed specifically to answer a set of questions and, therefore, potentially more effective in answering those questions. The population studied could be more appropriate and the measures of exposure and outcomes may have a higher validity than health care databases. Despite the strengths, they also have limitations. The participants volunteer in the study, for example patients or HCPs, and it may take a long time to accrue enough participants. The follow-up time could also be long, especially for outcomes such as cardiovascular diseases and malignancies. Because these studies are usually done for specific diseases and medicines, the data may not be suitable for other uses. They often require specific case report forms (CRFs), data cleaning and monitoring, which make them unsuitable for other research questions, even regarding the same drugs or diseases.

See case study A for an example of a study with a RWD with primary data collection in combination with a trial and an existing data source.

2.1.3. Federated systems

The availability of many different RWD sources presents a unique opportunity to perform the same study using different sources. Consistency of the results, or lack thereof, will help to understand better the research question being evaluated by potentially enlarging the sample size, including diverse patient populations, enriching health care data, and/or prolonging the study follow-up time.

There are two approaches to performing a study using multiple data sources: pooling the raw data together; or analysing the data separately and then combining the results using, for example, a meta-analysis (for more information on meta-analysis, see the CIOMS Working Group X report on Evidence Synthesis and Meta-Analysis for Drug Safety[17]). The former may be problematic, as data sources may originally have been built and developed with different purposes, formats, privacy laws, access requirements, and governance, making the pooling of the raw data potentially very difficult, if not impossible. The latter, called federated system, is more appropriate and there are currently several systems available.

Different RWD sources consist of data collected for different purposes and with different designs, using different formats, and using a variety of codes for diseases, conditions and medicinal products and devices. In a federated system, these different codes could be used in their original forms or harmonised and standardised into a single system. Such standardised systems can be referred to as common data models (CDMs), which have been used since the late 1990s in the US, including in the Vaccine Safety Datalink[18] and health maintenance organisation (HMO) Research Network.[19,20]

Starting in the late 2000s, CDM has also been used by the Observational Medical Outcomes Partnership (OMOP),[21] the Observational Health Data Sciences and Informatics (OHDSI) programme,[22] and the European Health Data & Evidence Network (EHDEN),[23] and the Data Analysis and Real World Interrogation Network (DARWIN EU).[24]

Another federated system using a CDM approach is the US FDA Sentinel System. This is an active safety surveillance system for US FDA-regulated medical products, using a distributed database of primarily electronic claims data collected for routine healthcare delivery. In the distributed data environment where participating data partners maintain physical and operational control over electronic data at their sites, data analytic codes are developed centrally and distributed to each data partner to execute against data stored in a CDM at each site.[25]

The advantage of using a CDM lies is that investigators can use multiple data sources together for the same purposes, using uniform codes and the same statistical methods, with broader patient populations and greater details in health care utilisations. However, the use of a CDM might result in a loss of information when converting data from individual data sources into a CDM by selecting or creating key variables for the CDMs.

With various established CDMs, there is a need to harmonise the CDMs to support research and analyses across multiple data networks. The enhanced data infrastructure provides the capacity to support evidence generation that can inform regulatory and clinical decision making.[26]

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP), a network coordinated by the EMA, is another federated system.[27] Different from the US FDA Sentinel, ENCePP is a network of investigators using different European RWD sources separately without a CDM, and maintaining the original codes and formats in each source.

See case study D and case study E for examples of studies using federated data systems.

2.1.4. Other sources

Other important RWD sources are SRSs such as the UK's Yellow Card scheme, US FDA Adverse Event Reporting System, the US FDA Vaccine Adverse Event Reporting System, WHO Vigibase, and EudraVigilance, the latter is a system for managing and analysing information on suspected adverse reactions to medicines, which have been authorised or are being studied in clinical trials in the European Economic Area. In addition, bio-pharmaceutical companies usually have their own SRS specific for their products. While some consider SRSs to not be ideal for informing causal inference, they have been an important source for signal detection since the 1960s[28,29] and will remain so for the foreseeable future.

A SRS consists of individual case reports that can be spontaneously reported or solicited. These can be reported by patients, HCPs and other reporters (such as those who become aware of the cases and then report them to the producers, market authorisation holders of the products, or regulators). Therefore, all observations reflect events, whereas the population (users of medicinal products) from which these events arise are not known. For this reason, the incidence of the events cannot be estimated without external data on the size of the exposed population. Another weakness is its cross-sectional nature, which means that there is usually no or only limited follow-up on individual patients, which is critical for evaluating associations between medicinal products and events. Other well-known weaknesses include under-reporting (not all events are reported), stimulated reporting (the reporting of events can be increased by factors like publicity), differential reporting (events related to certain medicines may be more likely to be reported than events related to others), and poor data quality in terms of low validity and completeness (e.g. the same event resulting in multiple reports).[30,31,32] Another limitation is the Weber effect, in which there is a gradual increase in reporting in early years after launch.[33]

Despite their limitations, SRSs play a key role in identifying and addressing safety issues. One of the SRS's strengths is the large amount of data that allows detection of rare events that cannot be identified from clinical trials and for detection of different signals simultaneously. For example, progressive multifocal leukoencephalopathy and phocomelia were first reported in SRSs, and such systems were proven useful in addressing the issues appropriately. Another strength is that SRSs are more frequently updated than other data sources.

Although SRSs have been used for signal detection for decades, given the limitations mentioned above, especially the lack of denominator (population at risk) information and follow-up, they are not suitable for signal assessment.

Cross-sectional survey databases, such the US National Health and Nutrition Examination Survey database, are other RWD sources that can play a key role in the evaluation of the burden or prevalence of diseases. The survey participants are usually representative of the population and thus the estimates of prevalence are generalisable to that population. A survey provides a snapshot of the population at a point in time or within a period of time, but given the lack of follow-up, they are not suitable for estimating risks. Moreover, many of the survey databases do not include information for a specific medicinal product and, therefore, cannot be used to evaluate the safety or benefit of a particular product.

2.2. Emerging data sources

The 21st Century Cures Act in the US and analogous initiatives elsewhere place additional focus on the potential for novel data sources to support active safety surveillance and regulatory decision making.

The introduction of modern computing, mobile devices and wearables, which may have biosensors or are used as input devices, has resulted in a large increase of data volume, data types, and data manipulation options. These new technologies enable tracking of patients' habits, activities, and health status and the use of such connected devices has especially increased among the chronically ill and the elderly. Even traditional medical devices such as glucose monitors are being connected to obtain data for real-time patient assessment or for reporting purposes in clinical trials.

At the same time, other important advances are converging, such as improved access to genomics data, the adoption of machine learning models for data analysis, and the move towards personalised medicine with biosensor data and cloud storage/computing potentiating these changes.

The existing and often incomplete diagnostic and procedure codes assigned for clinical or administrative purposes have been used for some time in secondary data analyses, but frequently lack rich and detailed clinical information. Secondary data use of a wide variety of ancillary data attached (or not) to an EHR is essential for improved safety signal detection, personalised medicine, impactful clinical research, reduced health care costs and population health management.

2.2.1. Biosensor data

This RWD source is of growing importance due to the rapid development in the digital field. It comprises wearables such as oxygen sensors, blood pressure monitors and electrocardiographic measuring equipment. The US FDA has cleared the Apple iWatch as a sensor to detect atrial fibrillation and other arrhythmias.[34] This allows more effective monitoring especially in a challenging time such as the lockdown during COVID-19 pandemic, for which the US FDA issued specific guidance on *Enforcement Policy for Non-Invasive Remote Monitoring Devices Used to Support Patient Monitoring During the Coronavirus Disease 2019 (COVID-19) Public Health Emergency: Guidance for Industry and Food and Drug Administration Staff*.[35]

An important characteristic of biosensor data is that measurements are longitudinal and tend to be either chronologically continuous or on a regular schedule. One example of a biosensor includes patch-based electrocardiogram + accelerometry for continuous measurement of heart rate, heart rate variability, or R-R (two successive electrocardiographic R-waves) intervals, that enables more complex post-processing analytics. Another example is wrist-based photoplethysmogram and accelerometry for continuous measurement of heart rate and physical activity.

An important challenge of biosensors is the need for GxP-validated devices, which permit processing of data that are device agnostic or more specifically built to accommodate different wearables as well as multiple types of data with variable sampling rates.

The use of such data to prospectively identify adverse events is promising. However, data storage, processes and analytics need to be developed to improve its use.

2.2.2. Patient experience data

Patient experience data (PED) considers perspectives, needs and priorities related to a disease or condition. PED includes information about patient preferences and PROs. Patient preferences enable consideration of the relative desirability or acceptability of alternatives or choices of interventions or endpoints, or alternative care pathways. PROs refer to data concerning the patient's health condition or status, from the patient's view rather than from the view of a HCP or test, and are especially useful for assessing how patients feel, function and survive. PROs typically consider QoL, severity of symptoms, degree of physical function, satisfaction with care, side effects, adherence to health interventions and perceived value of a health intervention.

While it is worth mentioning the emergence of patient-generated information on social media concerning adverse events, reasons for changing treatments, non-adherence, and QoL, it is important to note that such data, like SRSs, do not provide information on the population denominator.

2.2.3. Curated electronic health records combined with ancillary data

Using specific methodology and a common data model (CDM), curated EHRs can be combined with ancillary data to provide well-defined outcome and exposure information.

The volume and variety of health care-related data, included in the EHRs, have continued to grow in recent years and have been considerably enhanced by omics data and the proliferation of imaging and other complex reporting. However, secondary use of these additions to the EHR is currently challenging because the data tends to be stored in different disconnected systems and cannot be properly used without a curation process (the process of organising and managing RWD sources so they are readily available for access and use for different research purposes).

A comparison between several EHRs with and without curated data showed 67% concordance when relying on structured data alone versus 97.5% concordance among curated records.[36] Another challenge is that in some cases data codes are not uniform. Each data source has its own coding system and different ways of assigning codes to medicines are employed without following national or international standardisation.[37,38] For example, in oncology, we need to be clear about definitions such as overall survival, disease-free survival, objective response rate, complete response rate, and progression-free survival.

Even with curated data, different approaches can lead to different results. In oncology, such issues have been addressed using a specific CDM such as minimal Common Oncology Data Elements (mCODE).[39] Indeed, the wave of natural language processing approaches being added to easily implemented machine learning models may disrupt the manual curation process.

2.2.4. Text or images from radiology information systems as data

The field of radiology has been an early adopter of digital workflows and electronic integration and thus tends to have a more mature information system that virtually eliminates use of non-digital data. However, despite the existence of large amounts of digital data, secondary use of images and their associated reports has lagged due to lack of integration and appropriate methods. Effective use of this type of data requires personalised image interpretation (e.g. by a radiologist caring for the patient), discovery of new imaging markers, and wider data use by non-radiologists. However, such data are currently stored in complex and fragmented repositories under multiple layers of digital locks, which often precludes such uses.

An adverse event such as pulmonary embolism can be readily identified using computed tomography angiography (CTA), which is the test of choice. Nonetheless the actual positive predictive value is low (10%) due to the difficulty of selecting patients with a high pre-test probability. However, machine learning models using RWD on clinical, laboratory and other radiological information (e.g. chest x-ray), from large numbers of patients, could presumably be used to risk-stratify new patients and increase the CTA positive predictive value. The performance of these models (sensitivity, specificity, positive predictive value, negative predictive value, and accuracy) will have to be evaluated before implementation.

2.2.5. Output from laboratory information systems as data

Laboratory information systems are another rich source of secondary data that can be used for numerous purposes including adverse event identification and health outcomes research. Laboratory data can aid decision making and help to measure endpoints, outcomes, or exposures either alone or in algorithms, thus helping the data curation process.

Although laboratory data has been routinely used as secondary data for research, laboratory information systems is plagued by use of local, idiosyncratic and sometimes redundant and/or ambiguous names (or codes) rather than unique, well-organised codes from a standard ontology. As a result, secondary use of laboratory data requires investigators to invest considerable time cleaning the laboratory dataset. While efforts are in progress to improve laboratory information systems such as by using or mapping Logical Observation Identifiers Names and Codes (LOINC)[40] (see Chapter 1), coverage is not perfect ranging from 73 to 90% for a reference laboratory, which handles both common and specialised tests.[41]

In conclusion, the integration of specialised tests with images and unstructured text data is still for the future as the lack of standardisation has forced investigators to rely on one-off integration efforts.

2.2.6. Output from structured genomics investigation as data

Genomics emerged in the 1980s with the advent of efficient nucleic acid sequencing and was helped by the confluence of genetics, statistics, and large-scale datasets openly accessible to investigators.[42] The broad distribution of open datasets has required the creation of large-scale dataset repositories such as the European Nucleotide Archive (ENA), GenBank, National Center for Biotechnology Information (NCBI), Protein Data Base (PDB) , and Sequence Read Archive (SRA). Two consequences of these repositories have been the early adoption of a small set of standard data formats, and the open-source software frequently stored in GitHub, a platform for storing, managing and sharing software codes.[43]

The difficulty for an investigator is the need to combine genomics data with phenotype data. Few cohorts or registries with such merged data are available for analysis. One example is the Genetic Epidemiology Research on Aging (GERA), which involves 78 000 participants and 55 billion bits of genetic data, linked with comprehensive longitudinal EHRs as well as survey data on participant's health habits and background.

Merged phenotype–genotype databases provide a unique opportunity for advanced analytics on safety not only in clinical trials, but also for post-marketing studies. However, evaluation of medicine safety in the genomics space would need integration of a vast amount of continually changing data.

Two important issues cloud the bright future of pharmacogenomics: data ownership and privacy (see Chapter 4). The researcher's perspective is that open data would lead to better genotypes linked to phenotypes, while companies or even nations often seek ownership and control over large datasets given their obvious medical and commercial value.

The interest in the potential use of data from social media for safety surveillance has been increasing in the last decade. One study showed concordance of the numbers of adverse events from twenty SOCs involving medical products in Twitter and those reported in the US FDA Adverse Event Reporting System.[44] This concordance does not necessarily mean that social media data is reliable for signal detection, as the number of adverse events alone is not sufficient to define a signal.

A study under the European Innovative Medicine Initiative, IMI WEB-RADR (WEB-Recognising Adverse Drug Reactions) showed that 'broad-ranging statistical signal detection in Twitter and Facebook,

using currently available methods for adverse event recognition, performs poorly and cannot be recommended at the expense of other pharmacovigilance activities.'[45] Another study showed that, if the data were limited to patient groups, these signal detection methods performed better with sensitivity ranging from 29 to 50.6% and specificity from 86.1 to 95.5%.[46] This study also showed that up to 37.5% of the adverse events could have been detected earlier compared to the SRSs.

Despite its limited use for signal detection, social media data present a great potential for other purposes. For example, it could be used to evaluate the trends of the numbers of events reported while using medicines. While these numbers are not signals, they could be used to help prioritise events to evaluate further using more reliable data sources. In addition, it could also be used to evaluate the reasons for people rejecting vaccines or stopping medications.

Finally, as there has been a growing recognition of the importance of incorporating patients into decision making throughout the lifecycles of medicines and medical devices, social media is an important RWD source to obtain patient perspectives, including patient preferences and PROs.[47,48,49,50,51,52]

Chapter 2 – References

1 Mazzaglia G, Straus SM, Arlett P, da Silva D, Janssen H, Raine J, et al. Study design and evaluation of risk minimization measures: a review of studies submitted to the European Medicines Agency for cardiovascular, endocrinology, and metabolic drugs. *Drug Safety*. 2018;41(2):191–202. https://doi.org/10.1007/s40264-017-0604-4.

2 Santoro A, Genov G, Spooner A, Raine J, Arlett P. Promoting and protecting public health: how the European Union pharmacovigilance system works. *Drug Safety*. 2017;40(10):855-869. https://doi.org/10.1007/s40264-017-0572-8.

3 Hall GC. Validation of death and suicide recording on the THIN UK primary care database. *Pharmacoepidemiology and Drug Safety*. 2008;18(2):120-131. https://doi: 10.1002/pds.1686.

4 Smith Jervelund S, De Montgomery CJ. Nordic registry data: value, validity and future. *Scandinavian Journal of Public Health*. 2020;48(1):1-4. https://doi: 10.1177/1403494819898573.

5 Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990;1(1):43-46.

6 Noble W. How does multiple testing correction work? *Nature Biotechnology*. 2009;27(12):1135-1137. https://doi.org/10.1038/nbt1209-1135.

7 Robb MA, Racoosin JA, Sherman RE, Gross TP, Ball R, Reichman ME, et al. The US Food and Drug Administration's Sentinel Initiative: expanding the horizons of medical product safety. *Pharmacoepidemiology and Drug Safety*. 2012;1;21(1):9. https://doi:10.1002/pds.2311.

8 Woodcock J, Behrman RE, Dal Pan GJ. Role of postmarketing surveillance in contemporary medicine. *Annual Review of Medicine*. 2011;18;62:1-0. https://doi.org/10.1146/annurev-med-060309-164311.

9 Bate A, Hornbuckle K, Juhaeri J, Motsko SP, Reynolds RF. Hypothesis-free signal detection in healthcare databases: finding its value for pharmacovigilance. *Therapeutic Advances in Drug Safety*. 2019;5;10:1-9. https://doi.org/10.1177/2042098619864744.

10 Bate A, Hornbuckle K, Juhaeri J, Motsko SP, Reynolds RF. Hypothesis-free signal detection in healthcare databases: finding its value for pharmacovigilance. *Therapeutic Advances in Drug Safety*. 2019;5;(10):1-9. https://doi.org/10.1177/2042098619864744.

11 Gould AL. Generating and confirming hypotheses. *Pharmacoepidemiology and Drug Safety*. 2010;19(5):533-536. https://doi.org/10.1002/pds.1928.

12 Walker AM. Orthogonal predictions: follow-up questions for suggestive data. *Pharmacoepidemiology and Drug Safety*. 2010;19(5):529-32. https://doi.org/10.1002/pds.1929.

13 Wang SV, Kulldorff M, Glynn RJ, Gagne JJ, Pottegård A, Rothman KJ, et al. Reuse of data sources to evaluate drug safety signals: when is it appropriate? *Pharmacoepidemiology and Drug Safety*. 2018;27:567-569. https://doi:10.1002/pds.4442.

14 Skovlund E, Leufkens HG, Smyth JF. The use of real-world data in cancer drug development. *European Journal of Cancer*. 2018;1;101:69-76. https://doi.org/10.1016/j.ejca.2018.06.036.

15 Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clinical Pharmacology and Therapeutics*. 2019;106(1):36-39. https://doi:10.1002/cpt.1426.

16 The United States Food and Drug Administration (FDA). Guidance for industry: diabetes mellitus—evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes. (PDF accessed 15 September 2022).

17 Evidence Synthesis and Meta-Analysis: Report of CIOMS Working Group X. Geneva, Switzerland: Council for International Organizations of Medical Sciences (CIOMS), 2016. https://doi.org/10.56759/lela7055.

18 Chen RT, Glasser JW, Rhodes PH, Davis RL, et al. Vaccine Safety Datalink project: a new tool for improving vaccine safety monitoring in the United States. The Vaccine Safety Datalink Team. *Pediatrics*. 1997 Jun;99(6):765-73. https://doi.org/10.1542/peds.99.6.765.

19 Ross TR, Ng D, Brown JS, Pardee R, et al. The HMO Research Network Virtual Data Warehouse: A Public Data Model to Support Collaboration. *EGEMS* (Wash DC). 2014 Mar 24;2(1):1049. https://doi.org/10.13063/2327-9214.1049.

20 Platt R, Davis R, Finkelstein J, Go AS, et al. Multicenter epidemiologic and health services research on therapeutics in the HMO Research Network Center for Education and Research on Therapeutics. *Pharmacoepidemiol Drug Saf*. 2001 Aug-Sep;10(5):373-7. doi: 10.1002/pds.607.

21 Observational Health Data Sciences and Informatics (OHDSI). The Observational Medical Outcomes Partnership (OMOP). (Website accessed 15 September 2022).
22 Observational Health Data Sciences and Informatics (OHDSI). OMOP Common Data Model. (Website accessed 2 June 2023).
23 European Health Data & Evidence Network (EHDEN). (Website accessed 3 July 2023).
24 European Medicines Agency (EMA). Data Analysis and Real World Interrogation Network (DARWIN EU®). (Website accessed 11 March 2024).
25 Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiology and Drug Safety*. 2012;21:23-31. https://doi.org/10.1002/pds.2336.
26 The Office of the National Coordinator for Health Information Technology. Common Data Model Harmonization. (PDF accessed 15 September 2022).
27 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). What is ENCePP? (Website accessed 7 April 2023).
28 Almenoff J, Tonning JM, Gould AL, Szarfman A, Hauben M, Ouellet-Hellstrom R, et al. Perspectives on the use of data mining in pharmacovigilance. *Drug Safety*. 2005;28(11):981-1007. https://doi.org/10.2165/00002018-200528110-00002.
29 Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiology and Drug Safety*. 2009;18(6):427-436. https://doi.org/10.1002/pds.1742.
30 Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systematic review. *Drug Safety*. 2006;29(5):385-396. https://doi:10.2165/00002018-200629050-00003.
31 Dusetzina SB, Higashi AS, Dorsey ER, Conti R, Huskamp HA, Zhu S, et al. Impact of FDA drug risk communications on healthcare utilization and health behaviours: a systematic review. *Medical Care*. 2012;50(6): 466-478. https://doi:10.1097/MLR.0b013e318245a160.
32 Rawlins MD. Spontaneous reporting of adverse drug reactions. I: the data. *British Journal of Clinical Pharmacology*. 1988;26(1):1. https://doi:10.1111/j.1365-2125.1988.tb03356.x.
33 Weber J. Epidemiology of adverse reactions to nonsteroidal anti-inflammatory drugs. In: Rainsford KD, Velo GP, eds. Advances in Inflammation Research. Vol 6. New York, NY: Raven Press; 1984:1-7.
34 The United States Food and Drug Administration (FDA). Statement from FDA Commissioner Scott Gottlieb, M.D., and Center for Devices and Radiological Health Director Jeff Shuren, M.D., J.D., on agency efforts to work with tech industry to spur innovation in digital health. (Website accessed 29 September 2022).
35 The United States Food and Drug Administration (FDA). Enforcement Policy for Non-Invasive Remote Monitoring Devices Used to Support Patient Monitoring During the Coronavirus Disease 2019 (COVID-19) Public Health Emergency (Revised). Guidance for Industry and Food and Drug Administration Staff. (PDF accessed 29 September 2022).
36 Choi M, Song J, Bui CN, Ma E, Chai X, Yin L, et al. Costs per patient achieving remission with venetoclax-based combinations in newly diagnosed patients with acute myeloid leukemia ineligible for intensive induction chemotherapy. *Journal of Managed Care & Specialty Pharmacy*. 2022;28(9):980-988. https://DOI:10.1200/JCO.2020.38.29_suppl.307.
37 Leal LF, Osorio-de-Castro CGS, de Souza LJC, Ferre F, Mota DM, Ito M, Elseviers M, Lima EDC, Zimmernan IR, Fulone I, Carvalho-Soares MDL, Lopes LC. Data Sources for Drug Utilization Research in Brazil-DUR-BRA Study. *Frontiers of Pharmacology*. 2022;14;12:789872. https://doi:10.3389/fphar.2021.789872.
38 Justo N, Espinoza MA, Ratto B, Nicholson M, Rosselli D, Ovcinnikova O, García Martí S, et al. Real-world evidence in healthcare decision making: global trends and case studies from Latin America. *Value in Health*. 2019;22(6):739-749. https://doi:10.1016/j.jval.2019.01.014.
39 minimal Common Oncology Data Elements (mCODE). (Website accessed 15 September 2022).
40 Health Information Technology Standards Committee. Letter to: Dr. Farzad Mostashari, National Coordinator for Health Information. 2011. (PDF accessed 15 September 2022).
41 Lin MC, Vreeman DJ, McDonald CJ, Huff SM. A characterization of local LOINC mapping for laboratory tests in three large institutions. *Methods of Information in Medicine*. 2011;50(02): 105-114. https://:doi:10.3414/ME09-01-007.
42 O'Brien SJ. A decade of GigaScience: A perspective on conservation genetics. *GigaScience*. 2022;11,1-3. https://doi.org/10.1093/gigascience/giac055.
43 Github.com (Website accessed 29 September 2022).
44 Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. *Drug Safety*. 2014;37(5):343-350. https://doi.org/10.1007/s40264-014-0155-x.
45 Caster O, Dietrich J, Kürzinger ML, Lerch M, Maskell S, Norén GN, et al. Assessment of the utility of social media for broad-ranging statistical signal detection in pharmacovigilance: results from the WEB-RADR project. *Drug Safety*. 2018;41(12):1355-1369. https://doi.org/10.1007/s40264-018-0699-2.
46 Kürzinger ML, Schück S, Texier N, Abdellaoui R, Faviez C, Pouget J, et al. Web-based signal detection using medical forums data in france: comparative analysis. *Journal of Medical Internet Research*. 2018;20;20(11):e10466. https://doi:10.2196/10466.
47 European Medicines Agencies (EMA). The patients' voice in the evaluation of medicines. (PDF accessed 15 September 2022).
48 Abelson J, Wagner F, DeJean D, Boesveld S, Gauvin FP, Bean S, et al. Public and patient involvement in health technology assessment: a framework for action. *International Journal of Technology Assessment in Health Care*. 2016;32(4):256-264. https://doi:10.1017/S0266462316000362.
49 The United States Food and Drug Administration (FDA). Center for Biologics Evaluation and Research (CBER). (Website accessed 15 September 2022).
50 U.S. Department of Health and Human Services (HHS). (Website accessed 15 September 2022).
51 The United States Food and Drug Administration (FDA). Center for Devices and Radiological Health (CDRH). (Website accessed 15 September 2022).
52 van Overbeeke E, Whichello C, Janssens R, Veldwijk J, Cleemput I, Simoens et al. Factors and situations influencing the value of patient preference studies along the medical product lifecycle: a literature review. *Drug Discovery Today*. 2019;24(1):57-68. https://doi:10.1016/j.drudis.2018.09.015.

CHAPTER 3.
REAL-WORLD EVIDENCE FOR REGULATORY USE: KEY CONSIDERATIONS

The different sources of RWD (see Chapter 2), including primary and secondary data sources, derive from different health care-related settings. RWE can be derived from diverse data to answer different types of research questions and hypotheses about the effects of medical interventions' health outcomes.

The main interventions evaluated using RWE include prevention strategies, diagnostic methods, and treatments.

Traditionally, pharmacovigilance (the science and activities relating to the detection, assessment, understanding and prevention of adverse effects of medicines and vaccines) is an important research field where RWE is used to address safety issues. Evidence on safety of medicine is incomplete at medicines' approval because clinical trials are limited to their analysed patients' characteristics, sample size, and study duration; they may not adequately cover broad and complex exposure to concomitant medications, treatments, and food in real world. Therefore, post-marketing surveillance is necessary for immediate and long-term safety and effectiveness outcomes under real treatment conditions (e.g. considering incomplete medication adherence and the presence of comorbidities that may have been part of the exclusion criteria of the RCTs).

In each phase of the product lifecycle, including studies of effectiveness, research questions for the inclusion of RWE arise, as described in Chapter 1. The research question of interest is defined considering the evidence gaps, followed by the selection of an appropriate data source (primary or secondary data source), a study design and a protocol. Subsequently, the setting (i.e. primary or secondary data) of interest will be identified and the critical data source (e.g. administrative data or EHRs) will be determined.[1]

3.1. Assessing real-world data to generate evidence

It is important to have a good understanding of the type and the characteristics of the data source and to make the necessary evaluations and adjustments before using RWD for RWE generation, ensuring the data are of good provenance and fit-for-purpose for the research question. Differentiating between types of RWD studies, e.g. exploratory and hypothesis-testing studies according to the purpose of RWE generation, is recommended at planning. In this section, scientific, feasibility, and quality aspects are discussed with a focus on data source types.

3.1.1. Evaluating and selecting the database for specific purpose

Scientific fitness for purpose is critical for the selection of one or more databases. The attributes of a RWD source have to be suitable for the research question or study purpose, including the size

and representativeness of the study population and the availability of key variables about exposures, outcomes and covariates, including confounders with the method of data collection.

When considering the study population, it includes consideration of the entire population as well as subgroups. Confirmation that the selected data source covers the required participants for a sufficient duration to capture relevant exposures and outcomes for the planned study is essential. If the population covered is a subset of a data source, the variables to ascertain such a population should be available, and the number of patients should be sufficient for the study objective. A particularly important point of consideration is the representativeness of the patients in the data source and its impact on the external validity of the study findings to the real-world patient population of interest.

Especially in the secondary use of existing data, it is critical to establish whether the key variables (eligibility criteria, exposure, outcomes, and covariates) for answering the scientific questions of the study are recorded reliably in the selected data source. If the required variables are not available in the data source, one could investigate the possibility of additional/ad-hoc data collection. The availability of the additional data collection can also be particularly useful in pragmatic design clinical trials that utilise existing RWD as a data collection platform.

In the secondary use of existing data, the definition of the variable at the time of data collection needs to be investigated in detail. When using existing data collected over a long period, attention should be paid to changes in the definition of common variables in the relevant area, including disease classification. Absence of protocol-determined definition or target populations with detailed baseline description and standardisation of outcomes measures may critically impact the reliable interpretation of the findings.

A single database may not be sufficient for a given research question and multiple databases may need to be used. If that case, the principles of scientific considerations of fitness-for-purpose apply to each database.[2] When multiple data sources are used, proper work processes for data ingestion and harmonisation of datasets into a CDM are extremely important. If data linkage is needed, the right data elements for patient linking must be present.

Potential biases such as data availability bias and selection bias also need to be considered. Refer to case study B and case study C for utilisation of diverse databases in multiple countries with use of common standards and data models.

Depending on the time point of the medicine lifecycle, requirements may differ. Especially for initial market authorisation of a product, a favourable benefit–risk balance needs to be based on verifiable source data available in strict timelines.

In-depth description of study population and pre-specified study performance ultimately support causal inference according to outcome measures not necessarily performed in clinical practice in a sufficiently standardised manner. Both missing data/low return rate of e.g. questionnaires as well as reasons for discontinuation can be very informative and contribute to the benefit–risk considerations.

3.1.2. Feasibility considerations for evaluation and selection of the database

Time frame for data availability

For secondary use of existing data, the time lag between data collection (for example, the occurrence of events) and data availability may be a pitfall for the data utilisation plan. Each database has a different data collection schedule and data management plan. Data utilisation may require proposal

and approval process, and an agreement or contract may also be required for release of data. As a result, data may not be available according to the requested schedule. Communication with database holders should start in the early phase in study planning.

Access to data

The use of individual data may be limited due to ethical reasons (scope of patient consent) or regulatory reasons. Different levels of access may be available for individual data. For some databases, which allow secondary use of individual data, only strongly anonymised information is allowed for privacy protection (e.g. medical institution names cannot be provided and date data is converted to days). An agreement or contract may be required to identify data usage rights and scope of data usage.

In some RWD sources, only indirect access to data is possible. Under this model, data is analysed locally by data holders and third parties receive only aggregated data. For multi-RWD sources studies, this data will need to be combined at the aggregated level which limits some complex analysis.

Data feasibility framework

A data feasibility assessment framework called the Structured Process to Identify Fit-For-Purpose Data (SPIFD) is introduced to provide a systematic process for conducting feasibility assessments to determine if a data source is fit for decision making, helping ensure justification and transparency throughout study development, from articulation of a specific and meaningful research question to identification of fit-for-purpose data and study design.[3]

3.1.3. Quality considerations for evaluation and selection of the database

Data integrity

Data integrity refers to the completeness, consistency, and accuracy of data. Complete, consistent, and accurate data should be attributable, legible, contemporaneously recorded, original or a certified copy, and accurate. The system and procedure to maintain data integrity (for example, how accuracy and consistency are assessed and who is responsible) are of importance. For secondary use of existing data, the data governance of the data source should be evaluated.

Data governance is referred to as the total of activities, processes, roles, policies, and standards used to manage and control the data during the entire data lifecycle, while adhering to principles of data integrity. Security perspective is also required in the evaluation of databases. If any inadequacies are identified, evaluate the risks and investigate whether they can be adequately addressed.[4]

Transparency for data transformations including mixed data sources

Investigators should check if the data transformation/manipulation process is described in the lifecycle of the database or mixed data sources. Definitions to describe data transformation/manipulation as well as their processing need to be reviewed carefully. For mixed data sources, it is likely that a data transformation/manipulation process is required to perform analyses. In this case, a data transformation/manipulation dictionary and its process need to be defined and performed accordingly. This also ensures data traceability.

Data privacy rules of some countries require masking of small counts (< 5 or < 4), which limits the use of these data in analyses. This needs to be taken into consideration during the selection of RWD sources especially in rare indications or rare outcomes.

Regulations and good practice

Regulatory agencies generally require the sponsor to submit data from RCTs in support of regulatory review. However, the requirements for non-randomised studies are not entirely clear or consistent. For example, in the US, the submission of raw study data for regulatory submission of such studies seems important to address the concerns about the quality of both data and design in non-randomised RWE based on health care databases.[5,6] See Section 3.4 and Section 3.6 for more discussions on the topic.

The US FDA's draft guidance Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products requests that sponsors who submit non-interventional studies for regulatory review take responsibility for all activities related to the design, conduct, and oversight of the studies.[7] In the early stages of designing a non-interventional study intended for use in a marketing application, sponsors should ensure that they are able to submit patient-level data for the RWD.[8]

Data quality framework

Data quality frameworks such as Kahn's Framework,[9] recommendations from TEHDAS,[10] and Data Quality Framework for EU medicines regulation from EMA provide structured approaches to assess the quality of RWD. The recent EMA framework emphasizes several key dimensions of data quality that are crucial for evaluating the reliability and usability of RWD including reliability, extensiveness, coherence, timeliness, and relevance.[11]

3.2. Study design and methods

The choice of study design depends on the research question, type and availability of relevant data, and feasibility of the study. The selection of an appropriate design is important because it affects internal and external validity of the study results. The strengths and limitations of different study designs must be carefully considered to ensure the validity of the study results.

RWD/RWE can address research questions like estimating disease incidence/prevalence for orphan medicines, determining background adverse event rates, selecting comparators for single-arm studies, and providing descriptive evidence to support label changes. These data also explore treatment patterns, healthcare resource use, and patient-reported outcomes, enriching regulatory decision making beyond safety and efficacy assessments.

Emulating an RCT for designing studies using RWD is an approach that seeks to address the limitations of OSs in evaluating the safety and effectiveness of medical interventions. There are several advantages to conceptualising a non-randomised study using RWD as an emulated RCT. Most importantly, it clarifies thinking while making crucial design decisions such as inclusion criteria, duration of follow-up, and study endpoints, and reduces the potential for introducing error.

Emulating an RCT using RWD requires careful consideration of study design and data quality, as well as potential biases and confounding factors.[12]

3.2.1. Basic study designs of epidemiological observational research

The study designs of epidemiological studies are fundamental when using RWD to investigate the distribution and determinants of diseases in populations. Epidemiological OSs can be classified

into several basic designs as described below. Each of these designs has its own strengths and limitations, and the choice of design depends on the research question, type and availability of data, and feasibility of the study.

Cohort studies

Cohort studies follow a group of individuals over time to investigate the relationship between an exposure and a disease outcome. Prospective cohort studies follow individuals forward in time, collecting new data as time progresses. Retrospective cohort studies, on the other hand, use data that already exist at the time when the study commences.

Case-control studies

Case-control studies compare the exposure history of individuals with a condition (cases) to that of individuals without the condition (controls). For studies with primary data collection, this design is useful for studying rare diseases or when long-term follow-up is not feasible. Nested case-control studies are a variant of case-control studies within a larger enumerated cohort, where controls are selected from the same cohort as the cases. Population-based case-control studies are conducted on the entire population, and both cases and controls are selected from the same population at risk. Refer to case study F for nested case-control design.

Self-control case series studies

Self-control case series studies compare the occurrence of an event in during a period when an individual is exposed to a specific risk factor to the occurrence of the same event during periods when the individual is not exposed. This design includes only individuals who have experienced the study outcome and is useful for investigating the short-term effects of an exposure on a rare outcome.

Cross-sectional studies

Cross-sectional studies measure the prevalence of a disease and its associated risk factors at a particular point in time. These studies can provide information on the burden of disease in a population and help to identify risk factors for the disease. Because cross-sectional studies do not investigate whether the exposure came before the outcome or vice versa, cross-sectional associations generally provide limited evidence for causation.

Case series studies

Case series studies describe the clinical characteristics of a group of patients with a specific disease. These studies can provide valuable insights into the natural history of the disease and may generate hypotheses for further investigation.

3.2.2. Design elements and key considerations in their selection

Study populations

The successful implementation of a real-world study hinges on identifying the population that would most benefit from a given therapy or intervention. This is often achieved by anchoring the start of follow-up on an event that can affect subsequent treatment decisions, as one would do when designing a RCT. This can take the form of a new diagnosis, a laboratory value (e.g. an elevated glycosylated

haemoglobin in type 2 diabetes), an intervention (e.g. surgical procedure), or a prescription for a new medicine. Identifying a clinically relevant anchor point is critical as it establishes the temporality between potential confounders, the exposure, and the outcome. It is important to note that these considerations apply to both cohort and nested case-control designs where an underlying cohort has been identified and characterised.[13]

Historical controls differ from the contemporaneous controls in terms of their timing for cohort inception. For example, if an external control arm is constructed using RWD to support a single-arm clinical trial with a first patient enrolment in 2016, a historical control arm could be created using RWD collected before first patient enrolment in the clinical trial (i.e. before 2016).

In contrast, a contemporaneous control arm could be created if RWE was generated on or after the first patient was enrolled (e.g. using RWD collected in 2016 and onward). To account for any potential temporal changes — including changes in the SOC, medical practice or procedures, diagnostic criteria, and patients' beliefs and health behaviours — contemporaneous control cohorts are preferable to historical controls.

A particularly relevant potential update in medical practice is a change in who is eligible for treatment at all, which may drastically alter the severity of a disease in the patients included. However, there may be circumstances where the generation of external cohorts with contemporaneous data is not feasible, including the lack of availability of recent high-quality data, or scarcity of patients necessitating the use of historical data from multiple contiguous years. In these circumstances, the use of historical external controls may be acceptable under the condition that there were no large temporal shifts in the SOC, medical practice, patient management, or patient characteristics that are noteworthy.

Refer to case study A for the utilisation of natural history of disease data as external controls in comparison to single arm trial data.

Race and ethnicity

Constructs such as race and ethnicity merit additional care in the design and analysis of studies to generate RWE. Based on a recent review of studies conducted in the US and reported in major medical journals, the inclusion of race and ethnicity has increased over the past two decades but the quality of reporting has not.[14] Many health care databases contain limited if any data on race or ethnicity and lack critical details regarding the way those data were collected. The use of race or ethnicity and decisions regarding the representation of those who provide these data should be informed by an understanding of the community's interest in seeing themselves in the results while respecting privacy concerns.

Depending on the context, race or ethnicity may be regarded as a confounder and/or effect modifier. The use of race or ethnicity as a confounder should prompt an assessment of the role that historical and contemporary racism and its effects may play as important contributors even though data on those constructs may be less available[15].

Identifying the race or ethnicity of a person or group of participants, along with other sociodemographic variables, may provide information about participants included in a study and the potential generalisability of the results of a study and may identify important disparities and inequities. However, estimating differences in the effect of a treatment in specific populations based on race or ethnicity should only be done informed by an understanding of the local social, economic and institutional patterns that may influence health and health care.

Data sources with a representative sample of the population are likely to be underpowered to assess important differences, even if they are truly present, in the magnitude of the treatment effect across

subgroups. Interpreting differences found in safety or effectiveness as having a biological basis should only be done with robust evidence that other plausible explanations have been excluded.

Finally, as best practices are evolving in this area,[16] researchers are advised to seek up-to-date expert guidance on measurement, analysis and reporting of race or ethnicity.[17]

Outcomes definitions

Outcome definitions of RWD studies refer to the specific endpoints or measures that are used to evaluate the effectiveness or safety of a particular intervention or exposure in the study population. Selecting a clinical outcome measure in the real-world assessment of medicine effectiveness and safety involves careful consideration of disease or condition factors and data sources.[18] Outcome definitions adopted for this report are presented below, but other classifications of outcome definitions also exist.[19]

a. Clinical outcomes

These are outcomes that directly measure the health status of patients and are the most common category of outcome to be considered in RWD studies and are often related to specific diseases or conditions. The disease of interest may present with acute conditions, chronic conditions, transient or episodic conditions. Examples of clinical outcomes include morbidity, mortality, hospitalisation, symptom severity, disease-specific measures, and patients' functional measures.

Most clinical outcomes involve an objective assessment, most likely a diagnosis or assessment by a HCP. In real-world settings, these data are often recorded in a patient's medical record and may be coded as part of an EHR or administrative billing system using coding systems such as ICD-10 or ICD-11.

One needs to be cautious when defining outcomes for RWD studies, as clinical outcomes such as overall mortality defined as death from any cause may be more reliably recorded than outcome measures that are more subject to interpretation by individual HCPs such as depression or pain. Instruments such as diagnostic criteria, response criteria, and criteria for adverse events have been developed to help standardise the assessment of some conditions primarily used in clinical trials.

For RWD collected from a specific patient registry or a clinical study, the definitions of collected data should be thoroughly reviewed. RWD collected according to specific definitions can be an advantage when planning a RWD study. Refer to case study G for post-authorisation safety studies by using patient registry data with internationally harmonised survey items. Subjective assessment for clinical outcomes may also be considered for qualification for use in RWD studies.

Composite endpoints, with two or more component endpoints, are often used when the individual events included in the score are rare, and/or when it makes biological and clinical sense to group them. Using composite endpoints in RWD studies confers advantages by amalgamating multiple relevant outcomes. A recent RWD study using EHR emulating a randomised target trial to assess effectiveness of anti-diabetic medicines on preventing major cardiovascular events in patients with diabetes used a composite endpoint of stroke, myocardial infarction, and all-cause mortality. Composite endpoints may provide a comprehensive view of treatment efficacy, increasing statistical precision and reflecting broader clinical impacts.[20,21,22]

b. Patient-reported outcomes

A PRO is a measurement of any aspect of a patient's health status that comes directly from the patient without the interpretation of the patient's responses by a physician or anyone else. A PRO can

be measured by self-report or by interview provided that the interviewer records only the patient's response.[23] Examples of PROs include health-related quality of life (HRQoL), functional status, pain scores, satisfaction with treatment, and symptom burden.

HRQoL measures the impact of disease and treatment on patients' lives and is defined as 'the capacity to perform the usual daily activities for a person's age and major social role',[24] and often includes physical functioning, psychological well-being, and social role functioning.[18]

Many PRO questionnaires have been developed and validated. Examples of generic PRO questionnaires include Sickness Impact Profile (SIP, measurement of 12 domains and production of two subscale scores),[25] SF-36 (measurement of eight domains of physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional and mental health and production of component scores a physical component score, mental component score, and role/social component score),[26] and EQ-5D (measurement of scale in terms of five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression with generation of a single index score).[27]

Disease- or population-specific questionnaires, which may be more sensitive to symptoms affecting particular participants, are developed to detect differences and changes of QoL scores in response to disease or treatments.

Whenever possible, researchers should use PRO instruments that have been validated in the same kinds of people in whom the PRO will be used.

c. Surrogate outcomes

These outcomes are used as substitutes for clinical outcomes and are thought to predict clinical outcomes. Examples of surrogate outcomes include biomarkers, imaging findings, or laboratory tests considered to be associated with a particular disease or condition.

Surrogate endpoints may be used to reduce the follow-up period required to obtain results, thus are more commonly used in clinical trials than in OSs. However, if the surrogate outcome does not reliably predict the clinical endpoint of interest, unhelpful or even harmful interventions can look beneficial.[28] Therefore, before a surrogate endpoint can be accepted in place of a clinical outcome, extensive evidence of justified analytical and clinical validation must accumulate, including evidence from epidemiological studies and clinical trials to show that that the surrogate endpoint can reliably predict, or correlate with, clinical benefit in a context of use.

d. Economic outcomes

These are outcomes that measure the economic impact of an intervention or exposure. There are three cost components: direct costs related to health resource use, indirect costs related to productivity loss, and sometimes intangible costs (costs related to pain and suffering).

Direct costs correspond to the costs for patient care. They can be disentangled into two subgroups: direct medical costs and direct non-medical costs.[29]

The primary outcome of a cost-utility analysis is usually the cost per quality-adjusted life year (QALY), or incremental cost-effectiveness ratio (ICER). QALYs measure health as a combination of the duration of life and the HRQoL. ICER is calculated as the difference in the expected cost of two interventions, divided by the difference in the expected QALYs produced by the two interventions. The results of cost-utility analysis are compared with a threshold ICER; interventions with an ICER below this threshold are funded, those with an ICER above the threshold tend not to be. Inappropriate threshold values could cause new technologies to displace more health than they create.[30,31]

Exposure definitions

Selecting the appropriate exposure definition is critical in the real-world assessment of medicine effectiveness and safety. The three most common strategies are outline below, along with their strengths and weaknesses.

a. On-treatment exposure definition

The on-treatment exposure definition follows patients from the start until the end of their treatment. Thus, events of interest occurring during the follow-up period will be attributed to the patient's treatment at the time.[32] This exposure definition inherently assumes that the medicine has a reversible effect on the outcome (i.e. the effects disappear after treatment discontinuation).

This exposure definition is well adapted for acute outcomes thought to be prevented or caused while exposed to a given medicine (e.g. prevention of myocardial infarction and stroke). This definition answers the clinical question of what happens when patients are on the treatment.

Effectively implementing an on-treatment exposure definition requires two important assumptions. First, in the ideal setting, patients would refill a new prescription before the end of the previous prescription, thus ensuring uninterrupted use. While defining continuous exposure in this fashion would increase the specificity of the on-treatment exposure definition, it would reduce its sensitivity.[33] Indeed, this rigid definition does not account for small delays in refilling prescriptions, non-adherence, the pharmacokinetics of the drug, and the hypothesised effect on the outcome. Thus, on-treatment exposure definitions typically consider a grace period between non-overlapping successive prescriptions. The length of that grace period should be determined by the frequency of the prescribing (e.g. 30-day intervals), the pharmacokinetics of the drugs (e.g. drug half-life), and potential delays between outcome event onset and disease event detection or recording in the dataset. However, given uncertainties as to the optimal length of the grace period, sensitivity analyses should be conducted by varying the length of the grace period and assessing the impact on the effect estimates.

A second important assumption of the on-treatment exposure definition is that treatment discontinuation is unrelated to the outcome of interest. This assumption is not always appropriate, particularly if treatment discontinuation is related to disease progression (which is also associated with the outcome) or if the treatment was terminated because of prodromal symptoms of the outcome. In such situations, methods that account for potential informative censoring, such as inverse probability of censoring weighting, should be considered.[34,35,36]

b. As-started/intention-to-treat exposure definition

The as-started exposure definition, which is analogous to the intention-to-treat principle in RCTs, follows patients from the start of their treatment until the end of follow-up, regardless of treatment discontinuation.[37] This exposure definition answers the clinical question of whether to initiate a medicine versus another; it is about the intent of treatment.

For certain outcomes, the as-started exposure definition may be preferred over the on-treatment exposure definition. Indeed, the as-started exposure definition may be better suited for insidious outcomes with delayed onset, such as cancer incidence (especially if it is thought to have an irreversible effect on the medicine). However, the as-started exposure definition can be subject to important exposure misclassification, especially with prolonged follow-up. While this would generally lead to a dilution of the effect estimates, this is not always the case.[38]

c. Time-varying exposure definition

In the time-varying exposure definition, patients are followed from a cohort entry point and their exposure status is allowed to vary over time. Therefore, patients can contribute person-moments to different exposure categories during the follow-up period

This exposure definition reduces the possibility of immortal time bias,[39] while having the advantage of dealing with complex exposure patterns. Refer to case study D for Cox regression models with time-varying variables and with adjustment for potential influencing factors to calculate hazard ratios for risk of cancer. For example, using a time-varying exposure definition may make it easier to compare patients on a triple therapy to patients on dual therapy. However, implementing this exposure definition on large cohorts of patients can be computationally demanding.

One common aspect of study design is the recruitment of new users, or participants who have not previously been exposed to the treatment or intervention being studied. The concept of the new user design is described in the next sections, and is in keeping with conceptualisation of non-randomised RWD studies as emulations of a target RCT.

New-user vs. prevalent user definition in designs using an active comparator

One way of emulating a trial is to conduct a new-user, active comparator design.[40] This process typically involves identifying an exposure of interest and an active comparator. Both the exposure and active comparator should be used on new users, which avoids prevalent user bias and ensures that patient characteristics are measured before exposures, avoiding adjustment for factors affected by the choice of treatment.[41] This is typically achieved by selecting a washout period where patients are naïve to the exposures of interest.

It is important to note that the washout period implies that some patients may have been previously exposed to the intervention of interest, but not during the washout period (e.g. one year before cohort entry). While there is no clear consensus on an optimal washout period, it should consider whether the hypothesised association between the exposure and outcome is irreversible or reversible.

An irreversible effect implies that patients previously exposed to a medication would remain at risk, even after treatment discontinuation. This is typically assumed to be the case for outcomes such as cancer, where patients may remain at risk long after treatment discontinuation. In such situations, it would be preferable to anchor cohort entry on the first-ever treatment episode during the study period.

On the other hand, a reversible effect implies that the risk returns to baseline some time after treatment discontinuation. In such situations, there may be some flexibility in selecting a treatment episode that satisfies a minimum washout period.

An essential feature of the new-user, active comparator design is the selection of a comparator group. The comparator group serves two main functions. First, it can help reduce confounding by indication, which is a major threat to the internal validity of non-randomised real-world studies. Indeed, patients requiring a new exposure necessarily have clinical characteristics that would dictate a change in therapy. Thus, by selecting a clinically relevant comparator, it is possible to mitigate the effects of this bias at the design stage.

When possible, the active comparator be a medicine used at the same disease stage as the primary exposure. Comparing exposures given at different stages of the disease (e.g. a first-line treatment vs a last-line treatment) can introduce time-lag bias, a form of confounding by indication that would be difficult to control in statistical analyses.[42]

Second, the use of an active comparator facilitates clinical interpretation of the findings. This is especially important when contextualising the risks and benefits of specific therapies with others for which there is clinical equipoise.

Prevalent new-user design

While the new-user, active comparator design has become an important tool, it provides an answer to a specific question: should we initiate treatment with Exposure B or Exposure A? However, in some clinical situations the question is whether we should initiate Exposure B versus continuing with treatment A. This important question is often addressed in trials. The comparator group consists of no active treatment or SOC (such as in the cardiovascular outcome trials of novel diabetes medicines). In these settings, the comparator group is prevalent either by its non-use status or continuing the treatment received before randomisation.

There are also situations where many users of Exposure B have a history of Exposure A. This can be because of treatment guidelines or formulary restrictions recommending or limiting the use of Exposure B to patients in whom Exposure A failed. The prevalent new-user design was specifically designed to address these real-world situations.

As with the new-user, active comparator design, the prevalent new-user design also selects new users of the exposure of interest and an active comparator. However, the difference is that the latter group is not necessarily composed of new users.

Briefly, in the prevalent new-user design, new users of Exposure B who do not have a history of Exposure A are matched to new users of Exposure A who do not have a history of Exposure B (similar to the new-user, active comparator design). However, new users of Exposure B who have a history of Exposure A are matched to users of Exposure A provided they have a similar duration of use of Exposure A at the time of the switch. Thus, both new users of Exposure B and matched users of Exposure A have the same prevalence and duration of use of Exposure A.

Time-conditional propensity scores are used to control for the confounding associated with switching to Exposure B versus continuing treatment with Exposure A.[43] As the comparator group includes prevalent users, careful selection of variables is required to avoid including variables potentially in the causal pathway. This study design was used to assess the cardiovascular safety of aromatase inhibitors in women with oestrogen-positive breast cancer. It compared patients switching from tamoxifen to aromatase inhibitors with patients continuing treatment with tamoxifen.[44] An important consideration is that switching from tamoxifen to aromatase inhibitors is a common treatment strategy unrelated to disease progression. Indeed, sequential treatment with aromatase inhibitors followed by tamoxifen was investigated in several trials, and thus the prevalent new-user design emulated these trials.[45] This is distinct from another study using a new-user, active comparator design comparing new users of aromatase inhibitors with new users of tamoxifen;[46] that study assessed whether the upfront initiation of these medicines is associated with cardiovascular events.

Confounders

As noted above, it is often useful to consider a real-world study as emulating a target trial that one would like to conduct to answer a given question. Emulating a trial requires thinking about the cohort entry point for the exposures of interest to ensure that the treatment and comparison groups are comparable. While clinical trials can achieve this comparability by randomisation, RWD studies can achieve it by, among other approaches, addressing the issue of confounders.

Confounding is one of the biggest challenges in working with RWE and plays an even more significant role when comparing treatment effectiveness with safety. Confounding is present when the true association between exposure and the outcome is distorted by a third variable: the confounder.

A confounder can be defined as a variable for which control by design or analytical adjustment is required to obtain unbiased estimates of the effect of an exposure on the outcome under study.

Confounding factors, if not controlled, may result in:

- An observed association when no real association exists;
- No observed association when a true association does exist;
- An underestimate of the association (*negative confounding*);
- An overestimate of the association (*positive* confounding).

Confounding by indication

Confounding by indication, also known as channelling or confounding by severity, is often found in pharmacoepidemiological research. Confounding by indication occurs when the choice for treatment depends on (known or unrelated) patient characteristics that are associated with the outcome being studied, such as severity of disease. In general, the methods described in this section can be applied to confounding by indication. However, in effectiveness studies it is more challenging to correctly deal with confounding by indication given that the primary outcome is association between the treatment and outcome.

Time dependent confounding

Time-dependent confounding refers to confounders that change over time. In contrast to RCTs where researchers can actively minimise treatment non-adherence, loss-to-follow-up, and some intercurrent events, RWD studies are much more likely to face these challenges. If information of a confounder at different time points is available (such as body weight and laboratory values), the confounder can be addressed using analytical approaches including longitudinal targeted maximum likelihood estimation (L-TMLE), longitudinal inverse probability of treatment and censoring weighting (L-IPTCW), and marginal structural models.[47,48, 49, 50]

Descriptions on bias and unmeasured confounding are provided in more detail in Section 3.2.4. Statistical methods to improve comparability (e.g. matching and adjusted analysis) are discussed in Section 3.3.

3.2.3. Study design considerations in context of randomised controlled trials

Traditional Phase 3 RCTs have long served as the gold standard for evidence of clinical efficacy and safety of medical products to support regulatory approvals. RCTs can provide treatment effect estimates that are precise, with high internal validity to support a causal inference.

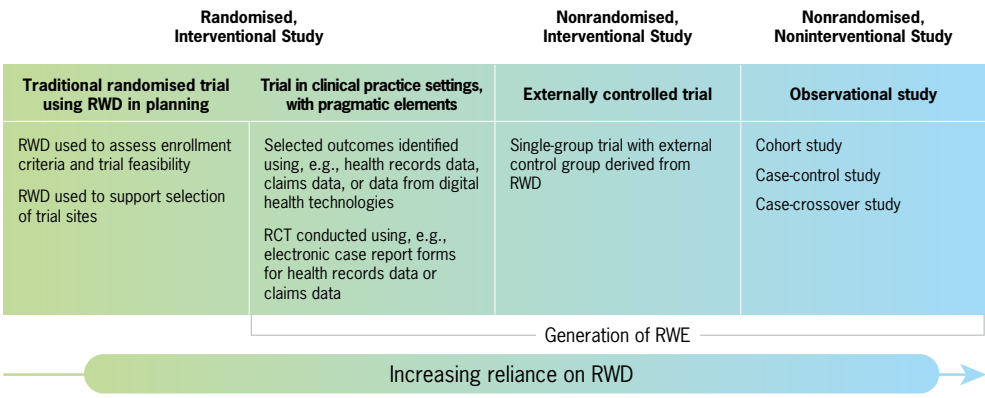
The reliability of RCTs is further supported by features that ensure accurate assessment of trial eligibility, treatment exposure (intervention) and outcomes. These features include a well-defined, specific trial entry and exclusion criteria, well-characterised, validated outcome measures, enhanced adherence to treatment and use of standardised study monitoring and capture of clinical outcomes that provide reliable and traceable data.

However, there are obvious limitations of traditional RCTs. They are resource intensive and slow to complete. Furthermore, they have limited generalisability (external validity) because the trials are too short in duration, trial participants are highly selected (may exclude older patients with comorbidities or concomitant medications) and sample sizes are small.

Figure 4 illustrates the various interventional and non-interventional study designs where RWD can be integrated into clinical evidence development of the effectiveness and safety of medical products, thus recognising that the degree of reliance on RWD varies with the type of study design. By definition, RCTs that incorporate RWD generate RWE.

Figure 4. Reliance on real-world data in representative types of study design

Source:[51] (Figure modified with permission)



Traditional randomised controlled trials using real-world data

Traditional RCTs are usually defined as an interventional research design in which one or more participants are prospectively assigned to one or more interventions including placebo to evaluate treatment effects on a health-related clinical, biological or behavioural outcome. Traditional RCTs are usually randomised, double-blinded, typically are supported by research infrastructure largely separate from routine clinical practice, and follow strict inclusion and exclusion criteria, protocol-defined standardised study monitoring and data collection procedures, including the use of detailed case report forms (CRFs) that are separate from routine medical records. This helps to ensure that high quality data with minimal variability are collected by specialised personnel.

Such traditional RCTs may integrate the collection of RWD elements outside of the research infrastructure to capture additional data that is relevant to the study. Routine EHRs, laboratory and pharmacy data may serve as useful sources of data. At times, these trials may rely on RWD from medical records for some clinical outcomes or need additional data for assessing relevant outcomes (radiographic or results of exercise stress tests). For example, in traditional RCTs of direct-acting oral anticoagulants against warfarin, double blinding limited the close monitoring of warfarin treatment to ensure that the international normalised ratio (a coagulation parameter) remained within the therapeutic range. This may have led to monitoring bias that impacted the adjudication of clinical outcomes. Integrating international normalised ratio monitoring from routine care could have helped investigators with outcome adjudication.

Interventional trials in clinical practice settings

a. Pragmatic randomised clinical trials

These trials are largely thought to answer important and relevant questions about the real-world effects of treatments in post-approval routine clinical practice settings.[52] The degree of pragmatism varies and such studies typically run on a continuum between traditional RCTs and observational non-randomised RWD studies.[53] They typically include a broader and more diverse study population of patients who are eligible to receive study interventions as part of routine clinical practice. Research participants are recruited from practice settings. Randomisation is usually at the provider or clinical practice level and not at the individual patient level. After randomisation, patients and providers make treatment decisions and no specific efforts are made to assure patient adherence to the intervention (such as a medicine) outside of routine practice. Adherence to treatment could be assessed through pharmacy claims/refills.

Primary and secondary outcomes could be collected from claims or EHRs or collected through limited electronic case report form (eCRF) with or without adjudication. While such trials can incorporate pragmatic elements, they can still have features to maintain rigorous standards for data collection.

1. Design considerations: Pragmatic trials are more suited to answering patient- and provider-relevant clinical questions related to comparative effectiveness and safety of medical interventions that are available and in use in routine clinical practice.
2. Study population and setting: Study population is usually composed of a broad and diverse patient population with a condition for which there are two or more approved interventions that are widely available in clinical practice. The study patients are recruited from routine clinical practice settings, usually community practices including general or speciality practices. The participating physician usually makes the study entry decision. Given pragmatic trials are embedded in routine practice and not randomised at the individual level, they may be conducted without an explicit patient consent with an approved waiver or may use a modified consent process (and should be discussed in more detail in Chapter 4).
3. Study Hypothesis, study treatment and comparator treatments: The primary hypothesis must be well defined, relevant and meaningful to participating physicians and patients. This study design is most appropriate when the goal is to demonstrate superiority of a study treatment against one or two available and accepted active treatment comparators on the selected study outcome(s). It is most suitable when the treatment effect difference between the treatment arm and the comparators in the selected primary study outcome is expected to be large. Treatment decisions follow routine practice as determined by the participating physician or practice treatment guideline. Interventions assessed in the study must be widely available and acceptable to participating practices and patients. The dosing and administration should ideally be uncomplicated and straightforward. Participating physicians may use protocol-defined regulatory approved treatments but may exercise greater flexibility in dose and regimen.
4. Outcome: The primary study outcome and secondary outcomes could be ascertained from practice EHRs and claims. Design considerations must take into account the following question to ensure accuracy and completeness of data collection. Can the investigator reliably capture the primary endpoint of interest from routinely collected data or require additional data collection using protocol defined eCRF? Can disease progression or changes be clinically assessed or require objective measures such as laboratory or imaging? Are there validated algorithms to identify and measure key outcomes? Can mobile technologies be used to fill in data gaps? Similar considerations apply to all other relevant outcomes such as emergency visits, hospitalisation, death etc.
5. Blinding: Usually patients and physicians are unblinded to treatments. Outcomes may be assessed and adjudicated in a blinded manner when possible. Randomisation at the practice level may help to assure initial balance in risk factors for the primary outcome event but may not mitigate

against variability due to selection and information biases such as selection of study patients, selection of co-interventions, degree of diagnostic intensity, reporting of outcomes and treatment discontinuation rates.

6. Adherence: Adherence to treatments could be assessed through pharmacy dispensing data (claims) and no special efforts to assure higher adherence are implemented. Without additional monitoring to ensure adherence to therapy, it is challenging to ensure comparability in adherence to treatment for drugs with a narrow therapeutic index such as warfarin (INR monitoring) when compared with novel agents that do not require INR monitoring.
7. Study Monitoring: The intensity and frequency of monitoring may range from routine practice procedures to limited additional protocol-defined requirements for follow-up as determined clinically appropriate by participating practice physicians. Safety monitoring and reporting may be streamlined to report serious adverse events and employ routine safety monitoring and reporting procedures of the clinical practice setting. The US FDA guidance on *Determining the extent of Safety Data Collection in Late-Stage Premarket and Post-approval Clinical Investigations*[54] is a useful reference to use.
8. Statistical analysis plan: Design and statistical analysis to address differences in baseline characteristics and impact of measured and unmeasured confounders is dealt with in Section 3.2.4. Pre-specification of the statistical analysis plan and inclusion of important prognostic and confounding variables in the data analysis is critical.
9. Limitations: There is a risk of falsely concluding that a treatment is more effective and safer than comparator treatments related to uncertainty of the robustness of evidence to support such a causal inference. Selection bias (patients not with target disease or difference in study outcome prognostic factors), information bias and other biases arising from lack of blinding and differential ascertainment of outcomes, study treatments, and co-interventions or concomitant medications can have a large impact limiting interpretability of study results. Additional limitations may arise from poor implementation of interventions, data quality and inadequate safety monitoring during the study.

b. Single-arm trials with external controls from real-world data

External controls, typically derived from past traditional RCTs, have been used as a control arm for single-arm trials. More recently, external controls derived from RWD are increasingly being used as controls for single-arm trials, especially for serious and rare diseases where an RCT is not feasible or where randomisation is highly unethical (e.g. for a promising treatment for a serious disease with a high unmet need).[55] Refer to case study A for utilisation of natural history of disease data as external controls in comparison to single arm trial data.

There were 433 single-arm trial submissions to HTA bodies between 2011 and 2019, and although this represented only 5% of the total submission volume over this period, single-arm trial submissions increased globally from eight in 2011 to 102 in 2019, a 13-fold increase.

Orphan status and the use of external comparators from RWD or prior trial are important determinants in acceptability of single-arm trial-based submissions. Data from registries, administrative EHR/claims and in some cases from case series or the literature have been used in such scenarios.

Use of external RWD control arms may pose important comparability challenges relative to the treatment arm due to systematic differences in the risk of study outcomes, outcome measure definitions and ascertainment methods, diagnostic procedures, medical practice, intensity of clinical monitoring, patient follow-up procedures, quality and completeness of data collection.

Regulations and guidance documents have indicated circumstances where historical control arm designs can be used. Codes of Federal Regulations 21CFR 314.126[56] indicates that historical

control designs are usually reserved for special circumstances. Examples include studies of diseases with high and predictable mortality (e.g. certain malignancies) and studies in which the effect of the medicine is self-evident (e.g. general anaesthetics).

ICH E10 (2001)[57] describes selection strategies for control groups in clinical trials intended to demonstrate efficacy. Section E suggests the inability to fully control for bias in external controlled studies except in situations where the effect of treatment is dramatic, and the usual course of the disease is highly predictable. US FDA has issued a guidance on externally controlled trials.[58]

Considerations of using external RWD control arm:

1. **Study patient population:** Use of external controls assumes similarity between trial patients and control group with respect to disease severity, duration of disease, prior treatments and important confounders that are prognostic of outcomes and the timing of the occurrence of outcomes. Differences in inclusion and exclusion criteria between patients from trial and from the external RWD control group may lead to selection bias and confounding limiting the validity of the inference from such studies. Design and statistical methods may be used to reduce bias. However, these important confounders (disease characteristics, current and prior treatments, important patient characteristics) may not have been assessed in the external RWD control group and the SOC may have changed over time.

Confounding by indication may be addressed to a certain extent by various study design elements to increase comparability of the trial and RWD control arms in important prognostic factors for the study outcome. These include techniques such as restriction, stratification, matching, modelling, and weighting. Sometimes matching on all the important variables may not be possible or efficient and the use of propensity score methods may be preferable. Selection bias could also be attributable to highly selected clinical trial participants. The consideration in comparability between RCT and external arm should come from both sides, including allowing RCT patients to better represent real-world patient population to minimise incomparability between two data sources.

2. **Primary and secondary outcomes/endpoints**: These should be well-defined objective endpoints, have similar definitions, ascertainment methods between the trial population and the external controls. Information bias arising from differences in the type of outcome measures, ascertainment method and timing of outcome assessment in the external RWD control arm relative to the trial patients may be a significant problem limiting inferences from such studies.

Information bias may arise from differences in the collection, recall, recording and processing of information. The problem may be compounded by differential missingness of data on important confounding variables (e.g. smoking). Information bias can also arise from non-differential (random) misclassification due to measurement errors in both the groups. Such non-differential information bias tends to underestimate treatment effect. On the other hand, differential information bias could work in either way, resulting in an overestimate or underestimate of the true treatment effect.

Epidemiologic strategies to avoid information bias include use of an appropriate study design, a well-designed protocol for data collection, handling and appropriate definition of exposures and outcomes.

3.2.4. Bias and unmeasured confounding

When using RWD, assessment of systematic error (bias) is critical for a study aimed at evaluating a medicinal product's treatment effect. However, we should acknowledge that the role of bias in epidemiology and pharmacoepidemiology has been described in many guidelines and reference works and therefore the aim of this paragraph should be to discuss the most important types of bias and

their relevance for our guideline and refer to other already existing guidelines for a more detailed description. For instance, the ENCePP Guide on Methodological Standards in Pharmacoepidemiology has a well-drafted chapter (4.2) on bias that is used as a starting document for this paragraph.[59]

Unmeasured confounding

A distinction can be drawn between measured and unmeasured confounding. An underlying assumption of RWD studies is that there is no unmeasured confounding. However, since no database contains information about all possible confounders, there will always be concern that one or more important unmeasured confounders exist, resulting in residual confounding when estimating treatment effect of a medical product. Therefore, in OSs and hybrid clinical trial designs, it is important to assess and question the potential impact of residual confounding.

It is important to distinguish the case when confounders are unknown from the case when confounders are known but not measured, as the methods to solve the issues may be different. Because the impact on final results can be significant, sensitivity analyses is strongly recommended to evaluate the robustness of study results against potential unmeasured confounders. See Section 3.3.7.

Selection bias

Selection bias relates to the selective recruitment of participants in a study who are not representative of the exposure (treatment) or the outcome in the population of interest. Examples are referral bias, self-selection bias, prevalence bias and protopathic bias.

Geriatric, paediatric or female patients should also be considered in the selection bias.

- Referral bias can occur if a patient is more likely to be recruited into a study due to certain patient characteristics and environment.[60] An example is when patients with a certain disease are referred to a tertiary or expertise centre in which they can receive certain specialised care. This may lead to a selection of certain patients, for instance more healthy patients who are easier to relocate.[61]
- Self-selection bias occurs when patients volunteer to enrol in a study because it is likely that their motivation for enrolling into the study makes them significantly different from the target population. For instance, if the internet is being used for surveys and health research self-selection bias may occur.[62] Alternatively, self-selection bias could occur when patients decide to drop out of a study for specific reasons, as opposed to randomly. This is why loss to follow up in a cohort study is a crucial aspect in determining the validity of that study.
- Prevalence bias is one in which the inclusion of prevalent users (for instance already using a treatment before start of follow-up) may introduce selection bias because they may be healthy survivors of the treatment. Others refer to prevalence-incidence bias or to Neyman bias.[63]
- Protopathic bias may relate to the issue of reverse causality. This can occur, for example, when a medicine is prescribed due to a headache while the headache itself was one of the early symptoms of some form of cancer. The study would show an association between the medicine and the cancer, even though the first symptom (headache) occurred before exposure to the medicine. This is described in more detail by Jessica Chubak and colleagues.[64]

Information bias

Information bias arises when incorrect information about either exposure or outcome or any covariates is collected in the study. It can be either non-differential when it occurs randomly across exposed/non-exposed participants, or differential when it is influenced by the disease or exposure status.

Examples of differential misclassification bias are recall bias (e.g. when control study cases and controls have different recall of their past exposures), and surveillance or detection bias.

- Missing data, or the lack of data/values in a data set, is a familiar problem that plays a role in all kinds of research and can contribute to information bias but may also lead to selection bias. The size of this problem is often larger in patient registries or health care databases than with RCTs or even carefully organised cohort studies, for several reasons. For example, it is unusual in registries for there to be any form of mandate to record data. Also, there are generally no periodic measurement moments. In addition, combining data from different data sources can increase the size of the missing data problem within a registry (for example, if there is unequal registration).
- Surveillance or detection bias arises when patients in one exposure group have a higher probability of having the study outcome detected, due to increased surveillance, screening or testing of the outcome itself, or of an associated symptom. For example, post-menopausal exposure to oestrogen is associated with increased risk of bleeding that can trigger screening for endometrial cancers, leading to a higher probability of detecting early-stage endometrial cancers. Any association between oestrogen exposure and endometrial cancer potentially overestimates risk because unexposed patients with sub-clinical cancers would have a lower probability of their cancer being diagnosed or recorded.[65] This may also occur in a study in which a new treatment is assessed in a single-arm trial and subsequently compared to historic controls (with no treatments).
- Immortal time bias refers to a period of cohort follow-up time during which death (or an outcome that determines end of follow-up) cannot occur. [66] Immortal time bias can arise when the period between cohort entry and date of first exposure to a drug, during which the event of interest has not occurred, is either misclassified or simply excluded and not accounted for in the analysis. Immortal time bias in OSs of drug effects[67] demonstrates how several OSs used a flawed approach to design and data analysis, leading to immortal time bias, which can generate an illusion of treatment effectiveness. This is frequently found in studies that compare groups of users against non-users.
- In many database studies, exposure status during hospitalisations is unknown. Exposure misclassification bias may occur with a direction depending on whether exposure to medicines prescribed before hospitalisation are continued or discontinued and if days of hospitalisation are considered as gaps of exposure, especially when several exposure categories are assigned, such as current, recent and past. The differential bias arising from the lack of information on (or lack of consideration of) hospitalisation during the observation period (called 'immeasurable time bias' in *Immeasurable time bias in OSs of drug effects on mortality*) can be particularly problematic when studying serious chronic diseases that require extensive medication and multiple hospitalisations.[68]

3.3. Statistical analysis of real-world data

3.3.1. Descriptive statistics and unadjusted analysis

Descriptive statistics are used to summarise and describe the basic features of the population, and can be used to assess imbalances between the study groups. These include measures of range, dispersion, and central tendency for continuous variables, number and percent for categorical variables, and plots for evaluating data distributions.[69] The standardised mean difference is often used to characterise the magnitude of differences in covariates between the exposure groups. The important first step in unadjusted analysis is to define a proper time scale and time origin for the data. A misspecification of the time origin can lead to biased estimates of all the outcome probabilities

of interest. The denominator of this estimated probability must include participants who are at risk and not participants without potential for experiencing the event at the time.[70]

Univariate or unadjusted analysis can be used to provide a preliminary assessment of which covariates are associated with exposure and/or study outcomes. Causal diagrams[71] are also an important tool for identifying the role that covariates play given our understanding of the temporal and causal relationships among these measures, the exposure, and outcomes of interest.

3.3.2. Estimation of absolute vs relative measures of effects

The reporting of relative effect estimates (e.g. hazard ratios, relative risks, and odds ratios) is routine and allows comparisons across settings with apparent ease. That said, relative measures can obscure potentially important differences when the background risk of the outcome varies between groups or settings. For example, when comparing a younger population with a low mortality rate (1/1000 person-years) to an older population with a higher mortality rate (100/1000 person-years), a constant relative effect of treatment (e.g. relative risk of 0.90) would lead to very different impacts of the intervention.

Estimates of absolute effects are valuable for weighing certain outcomes against others. For example, a large relative increase in the risk of a rare outcome (e.g. anaphylaxis) may be of less concern than a modest relative increase in the risk of a common outcome (e.g. myocardial infarction). Communicating the magnitude of relative effects improves if absolute effects (such as risk difference and number needed to treat) are included. Providing both absolute and relative measures of effect provides a range of stakeholders with more complete information on the potential benefits and harms of a given treatment.

The other elements of the study design and analysis will need to be informed by the choice of effect measures. For instance, some relative effect measures are unbiased when the outcome is assessed with perfect specificity (no false positives) and there are no differences by treatment group in the sensitivity. In contrast, the absolute effect measure (risk difference) is unbiased when the *sensitivity* is maximised, without differences by treatment group in the specificity.[72] Thus, the choice of effect measure has implications for selecting an outcome definition that maximises specificity or sensitivity.

3.3.3. Competing risk events

A competing risk is an event that precludes the outcome of interest from occurring for that individual. It is not merely the inability to observe the outcome of interest, but also eliminating the outcome from ever occurring, observed or unobserved. The most common competing risk is death. In any study in which mortality is not the outcome of interest, death before the event of interest will serve as a competing risk. Other competing risks are perhaps less obvious but equally important to address including, for example, hysterectomy in studies of uterine cancer, hospital discharge in studies of in-hospital mortality, complete mastectomy in studies of breast cancer recurrence.

Appropriate handling of competing risks is critical for the analysis plan. Many analyses erroneously treat competing risks like all other censoring events. This approach leads to the imputation of events for these individuals based on the observed event rate among those who remain uncensored in the analysis at later follow-up times. In doing so, the resulting estimates of the risk of the outcome of interest from the complement of the Kaplan-Meier curve will be inflated and therefore overestimate the risk.

If the competing risk is also of interest as an outcome relevant to the estimation of treatment effects, one simple approach is to create a composite outcome in which the occurrence of either outcome is used to estimate the time to event. For example, in a study designed to assess the effects of antiretroviral therapy among patients living with HIV, progression to AIDS or death can be used as a composite outcome rather than estimating the effect of treatment on progression to AIDS alone in which death would be a competing risk.

Statistical methods to handle competing risks include Fine-Gray subdistribution hazard model[73] and the Aalen-Johansen estimator[74] of the cumulative incidence of each event. Cumulative incidence probabilities can be estimated in consideration of competing risk events.[75] Group comparisons of the cumulative incidence probabilities over the entire interval can be tested by using Gray's test.[76] Log rank or weighted log rank test can also be used if the degree of competing risk occurrence can be deemed equivalent among the groups.

3.3.4. Adjusted analyses

Regression models are often used to estimate treatment effects adjusted or controlled for potential confounding variables. Refer to case study E for adjusted analysis by using Cox model.

Confounding variables are factors that are related to both the exposure of interest and the outcome of interest and not to the causal pathway from exposure to outcome. Variables that are potentially on the pathway are called intermediate variables and should not be controlled for, as controlling for them could affect the calculated effect of the exposure on the outcome.

Regression models are also often used in prognostic factor studies designed to determine patient, disease, and exposure/treatment characteristics, which influence clinical outcomes of the exposure/treatment.

Model assumptions and checking the model: The choice of regression model in RWD studies depends on the research question, the type of data, and the assumptions of the model. When applying regression modelling, careful attention must be paid to ensure that corresponding model assumptions are correct. For example, if Cox proportional hazards regression is used, then the proportional hazards assumption that the effects of the risk factors are constant over the follow-up time period, should be assessed.[77] If the validity of this assumption is questionable, then alternatives such as time-dependent covariates may need to be considered.

Interpretation of covariates: Variables should be handled and interpreted with care. For example, if the patient's age before treatment is entered as a continuous variable, the relative risk for every one-year increase in the patient's age is calculated. Another way of scoring the age effect would be to select a threshold. If the threshold is set for 50 years of age, and the value for patient's age under 50 years is 0 and over 50 years is 1 for the binary variable, the relative risk of the patient over 50 years of age with the patient under 50 years of age as a reference is calculated. Caution is needed when introducing a categorical variable with three or more non-ordinal values into the model. Creating dummy variables can be introduced to such variables.[78]

Presentation of results: The presentation of the results of the analysis should not be misleading and needs to be well considered. In the simplest studies this presentation usually involves a table of risk ratios for the variables of interest with appropriate confidence intervals and P values and a set of summary figures. Causal diagrams may be useful to understand the exposure and confounder effect estimates from a single model.[79]

3.3.5. Time-dependent covariates and time-varying effects

Most of the variables discussed in the previous sections are known when observation of the participants begins, or at the time origin. These are time-fixed covariates. Time-dependent or time-varying covariates are those whose value may change after the participant enters the study. Examples include continuous variables such as white blood cell or neutrophil count after starting chemotherapy, or binary variables indicating whether the patient developed febrile neutropenia after starting therapy or whether the patient is discharged by a given time. Refer to case study D for Cox regression models with time-varying variables for potential influencing factors to calculate hazard ratios for risk of cancer. Because the use of multi-variable models to adjust for variables observed during follow-up can introduce bias, alternative methods based on weighting should be used.[80]

3.3.6. Matching approaches for comparators focusing on propensity score-based methods

Matching is another approach to estimating treatment effects adjusted for potential confounding variables. With this approach, exposure groups are matched for the confounding variables at baseline. In this section, two widely used ways of matching are described: simple and propensity score matching.

In simple matching, the exposure groups are matched for the original confounding variables, such as gender, age, ethnicity, and comorbidities. In the propensity score matching, they are matched for the propensity score, which is the probability value that estimates the likelihood of receiving a certain treatment or exposure based on a set of observed covariates.

The use of the propensity score for matching to control for confounding was proposed by Rosenbaum and Rubin.[81] It is typically calculated by fitting a logistic regression model that predicts the probability of treatment assignment based on the covariates. Propensity scores can be used in sub-classification or stratification, matching, and weighting, and further adjusted using regression adjustment.[82,83,84,85] Refer to case study B for using propensity scores for matching.

Matching is primarily used when examining the effect of a point exposure that has two exposure levels, i.e. exposed and unexposed, to reduce the bias by reducing imbalance in the matched sample. The balance between the groups can be presented graphically or by comparing standardised differences across groups, which allows a reader to assess the balance in a similar manner to comparing randomised groups from a randomised clinical trial.

When using propensity score weighting, each individual's data is weighted by the inverse of their probability (IPTW) of the treatment they actually received to estimate the average treatment effect (ATE) in the total population. Alternatively, one can reweight only the comparator group to have the covariate distribution of patients who received the index treatment, which estimates the average treatment effect in the treated (ATT). Stratification by propensity score involves categorising individuals into groups based on their estimated likelihood of receiving a treatment, enabling more balanced and less biased comparisons in observational studies. These approaches aim to remove differences in covariate distribution between treatment groups and create similar groups where outcomes can be compared between treatment groups.

Matching can offer advantages over weighting with respect to robustness to assumptions about the exposure and outcome models and increased opportunities for customisation.[86] Matching has some costs as well, including generally less precision due to exclusion of unmatched observations. Although propensity score matching is an intuitive approach to causal inference, it tends to be more variable than other propensity score-based methods in a point treatment setting where there

is lost-to-follow-up.[87,88] In addition, matching on baseline covariates alone is inadequate when there is time-dependent confounding.

Propensity score-based methods (including propensity score matching) are unbiased when the propensity score is correctly modelled and there are no unmeasured confounders of the relationship between treatment and each of the outcomes under consideration. Again to note, these methods do not handle time-dependent confounding and thus only apply to estimating a marginal causal effect of a single time-point intervention on an outcome that is not subject to missingness or right-censoring, let alone, when the latter is informed by time-dependent covariates. Use of multiple analytic strategies as a sensitivity analysis (see next section) or doubly-robust estimators[89] may be useful, drawing strengths from both strategies.

3.3.7. Principles of sensitivity analysis

As described, the use of RWD comes with its own set of challenges, including potential bias and variability in the data, which can affect the reliability of the results. Sensitivity analysis is a series of analyses conducted to explore the robustness of inferences from the main estimator to deviations from its underlying causal assumptions and limitations in the data; thus, it can help address these issues for different scenarios, assumptions, and sources of variability.[90]

Analyses results are considered robust when they are consistent or unchanged by testing variations in underlying assumptions, although violations in assumptions that result in meaningful effect estimate changes provide insight into the validity of the inferences. Incorporating sensitivity analysis into RWD analysis for regulatory decision making can provide several benefits, including improved transparency and reliability of the analysis, increased confidence in the findings. Refer to case study H for sensitivity analyses to investigate the robustness of results.

Traditional sensitivity analysis is to test basic assumptions such as variable definitions and to consider the impact of unmeasured confounders. A study's underlying assumptions can be altered along several dimensions to evaluate robustness of results, including study definitions by modifying exposure, outcome and confounder definitions, and study design by changing or augmenting the data source or population under study, and modelling by modifying a variable's functional form or testing normality assumptions.[91,92,93,94]

Sensitivity analysis results can be presented in tables or graphs. Tables should allow readers to determine the influence of changes in assumptions. It is important to balance the benefits of including numerous sensitivity analysis results with the need for concise reporting.

3.3.8. Subgroup analysis

Subpopulations such as paediatric, geriatric, and racial/ethnic subgroups, or patients with comorbidities can be useful in sensitivity analysis to examine the robustness of study findings across different populations. It also can indicate the presence of effect measure modification, emphasising the need to acknowledge population heterogeneity in interpreting results. The analysis plan should specify whether effect measures will be estimated in such subpopulations to identify any effect measure modification. Refer to case study C for subgroup analyses stratified by age and sex subgroups.

3.3.9. Missing data

Incomplete data is a reality in all research but may be more extensive outside traditional randomised clinical trials. Missing data are defined as values that are not available and that would be meaningful for analysis if they were available.[95]

The extent to which data are missing and underlying dynamics that led to the missingness are important for determining how to handle these in the analysis. None of these methods will entirely make up for incomplete data collection, but the negative impacts can be mitigated to some degree. In-depth discussions of methods to address missing data are available elsewhere.[96] Recently, STRATOS (STRengthening Analytical Thinking for Observational Studies) initiative has published guidance framework for the treatment and reporting of missing data in OSs.[97]

Missing data are classified into three categories according to the reason for the data missing, and the degree of their relevance to the outcome: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). MAR is missing data that is related to the observed data but not to the missing data, and the value of the missing data that should have been obtained is considered to be explained by other observed data. MNAR is missing data related to the missing data and often depend on the observed data as well. The value of the missing data cannot be explained without data that should have been obtained.

There are several ways to approach missing data. It is important to highlight a common approach that is known to be inappropriate: complete case analysis. Excluding observations with missing values and only analysing those individuals who have no missing data is rarely acceptable due to the selection bias that results from conditioning on complete data.

Imputations are methods to supplement missing data values from other observed data. A last-observation-carried-forward, a baseline-observation-carried-forward, a mean value imputation, a random imputation method are examples of single imputation method. Multiple imputation addresses missing data by using other information about the individuals with missing data to impute the expected value for the missing information. For example, if data on body mass index (BMI) are missing for 10% of the study population, a predictive model, fitted to those with non-missing BMI, can be used to estimate the likely value for BMI for individuals where it is missing conditional on their age, sex, etc.

To account for the uncertainty introduced by imputing some values, multiple imputed datasets are created, analysed, and then the results are combined using Rubin's Rule in order to reflect the wider confidence intervals due to the imputation. For this method to be useful, it is necessary to be able to fit a reasonably good predictive model for the missing variable using information from the other available covariates including the outcome. Thus, it is more important to have a reasonable number of observations in which to develop this model rather than a given percentage of the data which is non-missing. For instance, a very large study with 100 000 observations may have 90% of the data on BMI missing and still be able to fit a predictive model within the 10% (n = 10 000) observations who are non-missing.

Statistical models are often used with imputation methods. Statistical models such as inverse probability weighting, mixed model for repeated measure, and pattern mixture model are often used in conjunction with imputation methods. Conventional statistical analysis of missing data has mainly used methods based on the MAR assumption using multiple imputation methods.

The recent Treatment and Reporting of Missing data in Observational Studies (TARMOS) framework[98] discusses the need for sensitivity analyses under the assumption that MAR is not valid. Refer to case study H for consultation with health authorities for the treatment of missing data for primary endpoint. The initial analysis plan was changed to perform complete case analysis. As previously noted, sensitivity analyses investigated the robustness of the results.

3.4. Evidence-generation, study registration, reporting, documentation and communication

Regulatory decisions impacting public health in the form of marketing authorisation approvals, and to some extent reimbursement decisions, have traditionally been based on clinical trials for which rigorous criteria to ensure data integrity have been developed. This includes: registration of protocols; pre-specifying analysis, masking study interventions from participants, investigators, endpoint adjudicators and analysts; publication; and results disclosure.

Similarly, trust in RWE by regulatory bodies will be engendered and their acceptance increased if generally accepted criteria for transparency are complied with.

Recent regulatory approvals based on RWE created an urgency to develop generally accepted processes that promote trust in the evidence-generation process. Transparency of the research process to enable decision makers to evaluate the quality of the methods used and the applicability of the evidence that results from the RWD studies will be key in this process.

Registration of RWD studies — particularly for hypothesis evaluating treatment effectiveness (HETE) studies — has been proposed to improve transparency, trust, and research replicability. Although registration would not guarantee better conducted RWD studies, it would encourage prospective disclosure of study plans, timing, and rationale for modifications.

While the focus of sponsors may be regulatory acceptance, other key stakeholders and decision makers, including patients, HCPs, learning health systems, and policy makers interested in bioethical and regulatory issues, will benefit from best practice standards.

To that end, several international professional societies including Duke-Margolis Center for Health Policy, ISPE, and ISPOR have issued recommendations.

A joint task force of ISPE and ISPOR recommended that investigators pre-register their RWE studies and post their study protocols on a public platform before starting studies to reduce publication bias and improve the transparency of research methods. Recognising the structural and practical challenges, the Real-World Evidence Transparency Initiative has outlined a pathway to move forward.[99]

RWE studies range from exploratory, hypothesis-generating study to HETE. Although exploratory analyses of secondary data are often necessary to understand the relevance and quality of the data for the proposed analysis, a concern is that analysts could make decisions on study design after seeing the preliminary results.

Without transparent pre-specification of hypotheses, data sources, protocols, and analysis plans, concerns about results-driven selection of study parameters and selective reporting on favourable findings can undermine confidence in the reported results of HETE studies. Thus, criteria for HETE are proposed to ensure specifically transparency and trust.[100]

The formulated general principles highlight the need to prospectively define study methods in evidence generation, registration, stakeholder alignment with regulatory authorities and HTA bodies before doing the study and transparent reporting. Another aspect is the ability to create audit trails (auditing the vendor, the database, the sponsor).

Applying the outlined principles to the extent possible for exploratory studies could improve transparency and trust into other designs as well and, therefore, be viewed as general recommendations.

Box 1. ISPE/ISPOR taskforce recommendations for HETE

Source:[101] (Figure reproduced with permission)

1. A priori, determine and declare that a study is a hypothesis evaluation treatment effectiveness (HETE) study or an exploratory study based on conditions outlined below.
2. Post a HETE study protocol and analysis plan on a public study registration site before conducting the study analysis.
3. Publish HETE study results with attention to conformance and/or deviation from the study protocol and original analysis plan. Possible publication sites include a medical journal, or a public web-site.
4. Enable opportunities to replicate HETE studies (i.e. for other researchers to be able to reproduce the findings using the same data set and analytical approach). The ISPE companion paper lists information that should be reported to make the operational and design decisions behind a RWD study transparent enough for other researchers to reproduce the conduct of the study.
5. Perform HETE studies on a different data source and population than the one used to generate the hypotheses to be tested unless it is not feasible (e.g. another data set is not available).
6. Authors of the original study should publicly address methodological criticisms of their study once it is published.
7. Include key stakeholders (patients, caregivers, clinicians, clinical administrators, HTA bodies, payers, regulators, manufacturers) in designing, conducting, and disseminating HETE studies.

Existing study registries (e.g. the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance Post-Authorisation Study (EU-PAS) register) are used for registering post-authorisation safety studies (PASS) by sponsors or research commissioned by regulatory bodies such as the EMA. ClinicalTrials.gov focusses on studies that collect primary data and lack many of the features needed for a study registry designed to improve transparency. Presently, sponsors disclose OSs with medicines, biologicals and vaccines, including over-the-counter products following internal company standards and recommendations of trade organisations of pharmaceutical manufacturers. Also, guidelines for Good Pharmacoepidemiology Practices (GPP) may apply.

For PASS studies, additional registration and results disclosure is required on the EU-PAS register.

Previous proposals have called for the registration of non-interventional studies[102,103,104] but the systems used and incentives to systematically register all studies have been unsatisfactory so far. It is hoped that this will improve with further collaborative efforts, such as the RWE Transparency Initiative, initially led by a partnership among ISPOR, ISPE, the National Pharmaceutical Council, and the Duke-Margolis Center for Health Policy. The long-term goal of this initiative is to make registration of HETE RWE studies routine in the way that the registration of clinical trials has become routine. In scope are particularly studies whose findings are intended to support decisions by regulatory agencies, payers, or other health care decision makers, including clinicians and editors of peer-reviewed journals who must decide whether or not to publish a HETE study.

The Real-World Evidence Transparency Initiative has identified practical steps to building on the foundation of existing study registries, identified issues that affect the practicality of the registration process, and considered how to facilitate routine registration of HETE RWE studies. Appropriate balance between the amount of detail registered and confidentiality required is critical for ensuring appropriate usage of the registry. For example, concerns about intellectual property rights in a

public registration may be addressed by temporary restriction of information to privileged users such as regulatory authorities.

Registration may also facilitate overcoming the concern about publication that is present in clinical trials, but even more so in RWE. The totality of evidence on a given topic requires that information about most studies on the topic, including from studies with negative results, be available to users. It is essential to compare study results and methods for a given hypothesis, including replications of studies.

The recommendation from the Joint ISPE/ISPOR group is to register each RWE study protocol, including key study parameters in a registry. The use of structured reporting templates to improve the readability of posted information is encouraged. Registered study protocols should be date stamped, including date-stamping of all revisions to the protocol with a rationale for each change.

Of particular importance is the requirement to pre-specify the analysis as it will address a number of broader issues such as:

- Blinding to protect the analysis;
- Selection of individuals from inserting bias;
- Specific concerns in external control arms;
- Use of blinding to outcomes to ensure that those fitting exposure (propensity score) model are blinded to the outcomes.

ISPE/ISPOR notes that in RWE, terminology for the same concepts varies around the world. Agreeing on terminology and which parameters from a large catalogue are the most essential to report for replicable research would improve transparency and facilitate assessment of validity.[105]

3.5. Reproducibility of real-world data studies

Reproducibility is the cornerstone of the scientific method. However, there have been concerns about the reproducibility of research in many scientific fields, including cancer biology,[106] psychology,[107] and economics,[108] as well as clinical trials.[109]

There have been several efforts to evaluate the replicability of studies in various disciplines, with results suggesting room for improvement. Recently, there was a systematic attempt to measure how we are doing in terms of the reproducibility of RWD studies.[110] This project identified a systematic, random sample of RWD studies published in leading medical and epidemiology journals; it then attempted to reproduce them using the same years of data from the same data sources, and the same methods as reported by the original investigators. The findings indicated that while most studies could be closely reproduced, a substantial minority could not. A few areas that contributed to difficulty with reproduction included 1) incomplete information on details of key design parameters (particularly temporality and code algorithms); 2) incomplete information about data version; and 3) internally inconsistent information between the text, attrition tables, design diagrams, and supplemental materials.

Reproducibility is closely related to how clearly scientific processes are communicated. When the steps taken to implement a study are unambiguous, we are better able to understand how the evidence was generated, to evaluate the validity of the methods, and to understand the reasons for apparent divergence from studies that seem to be asking the same question.

There are many different types of reproducibility. In the context of database studies, these include computational reproducibility, independent reproducibility, and conceptual replicability (or robustness).

Computational reproducibility is the ability to re-run the same code on the same data and get the same results. However, without clear natural language description about what scientific decisions are being implemented, it can be difficult for reviewers or decision makers to assess the validity and relevance of those decisions for the question of interest.

Independent reproducibility involves the ability to independently recreate the analytic cohort and analysis from the source data warehouse. This is an important type of reproducibility to have because it requires unambiguous reporting of design and implementation decisions. This level of clarity about scientific decisions facilitates assessment of their validity and relevance.

Conceptual replicability or robustness evaluation is about trying to address the same question or causal estimand using different data or methods.

Each type of reproducibility could be facilitated through the use of structured protocol templates like HARPER,[111] registering protocols, sharing code, and providing sufficient information on data sources.

3.6. Agreement between real-world data studies and randomised controlled trials

As previously noted, RCTs are considered the gold standard for evaluating the efficacy of medical products. RWE can provide valuable complementary evidence of drug effects under clinical practice conditions, and in populations in whom RCTs cannot be ethically conducted; however, there remain concerns about the credibility of RWE to support causal inference.

Decision makers are most concerned about bias when it comes to non-randomised, non-interventional studies. A natural benchmark for evaluating the validity of the causal inferences drawn from RWD studies is the concordance of the RWD study results with the results of an RCT.

There have been numerous one-off studies that compared results between published RCTs and RWD studies, with mixed results.[112] The credibility of RWD studies has suffered from apparent divergence in results between database RWE and trials. The RCT-DUPLICATE Initiative has a large-scale series of projects aimed at understanding when and how RWD studies can generate valid results and inform regulatory decision making.[113] Over 30 trials were systematically sampled from a variety of clinical areas and emulated using RWD. Some of the main takeaways from this project included:[114]

a. Simple measures of agreement in results between RCTs and RWD studies lack nuance and will not tell the whole story. When emulating an actual trial instead of a hypothetical trial, there will be design emulation differences *in addition to* potential biases. Researchers and reviewers often have to dig deep to outline, understand, and tease these apart.
b. Residual bias or random error are always potential explanations for observed divergence in results between a trial and a RWD study. However, when the divergence is driven by design emulation differences, the database study could be accurately targeting a different effect (for a different research question) than the trial.
c. Given low adherence in clinical practice, it can be challenging to replicate trial findings for outcomes with a long induction window or time varying hazard over extended follow up. Related to this point, in clinical practice, patients may not experience the benefit identified in trials that create ideal but unrealistic conditions to maximise their ability to detect an effect.
d. Comparisons of RCT and RWD studies typically use the result of a single trial as a reference standard. This does not take into account the uncertain replicability of a trial's findings even by other trials (which can go beyond chance).

Although the overlap in research questions that could be addressed with both RCT and RWD studies is limited, RCT-DUPLICATE[115] and other similar RCT emulation projects (Observational Patient Evidence for Regulatory Approval and uNderstanding Disease (OPERAND),[116] Center of Excellence in Regulatory Science and Innovation (CERSI)[117]) have demonstrated that when the data and design are fit for purpose, non-randomised database studies can come to similar conclusions about drug effects as randomised trials.[118,119]

However, the real benefit of non-randomised, non-interventional RWD studies is in how they can complement the evidence from RCTs. So, when considering which tool would be most appropriate in a given situation, an important point to consider would be: would the hypothetical target trial, that would address the need of the end user, provide evidence of the drug effects under ideal conditions or clinical practice conditions?

3.7. Quality tools for real-world data studies and real-world evidence reporting

Various tools exist to assess the completeness of reporting such as Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and quality of non-randomised studies such as Grading of Recommendations, Assessment, Development and Evaluations (GRADE). STROBE provides a checklist of items that should be described in any reports of OSs.[120] For example, STROBE advises that for data sources, each variable of interest, the source of the information, and detailed methods of measurement including diagnostic criteria, if applicable, should be provided.

Reporting of studies Conducted using Observational Routinely collected health Data for PhamacoepidEmiology (RECORD-PE) statement is extended from the REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement to include reporting guidelines specific to pharmacoepidemiological research. The RECORD-PE checklist is provided with the statement with intent for the improvement of the standards of reporting of pharmacoepidemiological research undertaken using routinely collected data.[121,122]

The GRADE methodology provides a transparent framework for developing and presenting summaries of evidence assessed per patient relevant outcome with regard to the respective certainty of evidence.[123,124,125] GRADE has four levels of quality of evidence (very low, low, moderate, and high).[126] Evidence from RCTs starts at high quality and evidence from observational data starts at low quality. The certainty in the evidence is increased or decreased depending on more detailed features of the studies in comparison to the question of interest. Categories of outcome specific assessment are risk of bias (see Section 3.2.4), indirectness, imprecision, inconsistency and risk of publication bias.

Using a specific checklist for observational studies, experimental (f.e. RCT) and observational (f.e. RWD) data can be integrated in one evidence synthesis.[127,128] GRADE provides further a systematic approach for developing recommendations in terms of an "evidence to decision framework".[129] It has been officially endorsed by over 100 organisations.

The Registry Evaluation and Quality Standards Tool (REQueST) is a recently developed tool to support HTA organisations and other actors in guiding and evaluating registries for effective usage for decision making by regulatory and HTA bodies.[130,131]

Chapter 3 – References

1 A Framework for Regulatory Use of Real-World Evidence. September 13, 2017. Duke Margolis Center for Health Policy (Website accessed 2 June 2023).
2 Gatto NM, Campbell UB, Rubinstein E, Jaksa A et al. The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework. *Clin Pharmacol Ther*. 2022 Jan;111(1):122-134. https://doi.org/10.1002/cpt.2466.
3 Gatto NM, Campbell UB, Rubinstein E, Jaksa A, et al. The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework. *Clin Pharmacol Ther*. 2022 Jan;111(1):122-134. https://doi.org/10.1002/cpt.2466. Epub 2021 Dec 1.
4 European Medicines Agency. Guideline on computerised systems and electronic data in clinical trials. Good Clinical Practice Inspectors Working Group (GCP IWG). 9 March 2023, EMA/INS/GCP/112288/2023.
5 Franklin JM, Platt R, Dreyer NA, London AJ, Simon GE, Watanabe JH, et al. When can nonrandomized studies support valid inference regarding effectiveness or safety of new medical treatments? *Clinical Pharmacology & Therapeutics*. 2022;111(1):108-115. https://doi.org/10.1002/cpt.2255.
6 Wang SV, Sreedhara SK, Schneeweiss S. Reproducibility of real-world evidence studies using clinical practice data to inform regulatory and coverage decisions. *Nature Communications*. 2022,31;13(1):5126. https://doi.org/10.1038/s41467-022-32310-3.
7 The United Stated Food and Drug Administration (FDA). Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. Guidance Document. (PDF accessed 15 October 2022).
8 The United Stated Food and Drug Administration (FDA). Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. Guidance Document. (PDF accessed 15 October 2022).
9 Kahn MG, Callahan TJ, Barnard J, Bauck AE, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS* (Wash DC). 2016 Sep 11;4(1):1244. https://doi.org/10.13063/2327-9214.1244.
10 Towards European Health Data Space (TEHDAS). TEHDAS develops data quality recommendations. 18.05.2022. (Website accessed 21 April 2024).
11 European Medicines Agency. Data Quality Framework for EU medicines regulation. EMA/326985/2023. 30 October 2023. (PDF accessed 21 April 2024).
12 Miguel A Hernán and James M Robins. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology*, Volume 183, Issue 8, 15 April 2016, Pages 758–764, https://doi.org/10.1093/aje/kwv254.
13 Suissa S. Brief Report: The Quasi-cohort Approach in Pharmacoepidemiology: Upgrading the Nested Case–Control. *Epidemiology*.2015;26(2):242-246. https://doi:10.1097/ede.0000000000000221.
14 Martinez RAM, Wilbur RE, Andrabi N, Goodwin AN, Smith NR, Zivich PN. Conceptualization, operationalization, and utilization of race and ethnicity in major medical journals. *medRxiv* The preprint server for health sciences. doi: https://doi.org/10.1101/2022.03.07.22271661.
15 Howe CJ, Bailey ZD, Raifman JR, Jackson JW. Recommendations for Using Causal Diagrams to Study Racial Health Disparities. *Am J Epidemiol.* 2022 Nov 19;191(12):1981-1989. https://doi.org/10.1093/aje/kwac140.
16 Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards. (Website accessed 4 June 2023).
17 Flanagin A, Frey T, Christiansen SL, AMA Manual of Style Committee. Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA*. 2021;326(7):621–627. doi:10.1001/jama.2021.13304.
18 Velentgas P, Dreyer N, Noujah P, Smith SR, and Torchia MM (eds.). Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Agency for Health Care Research and Quality (U.S.) Government Printing Office; 2013.
19 FDA-NIH Biomarker Working Group. Silver Spring (MD): Food and Drug Administration (US); Bethesda (MD): National Institutes of Health (US); BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Glossary. Published 28 January 2016. Last update 29 November 2021. https://www.ncbi.nlm.nih.gov/books/NBK338448/#IX-C.
20 Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. *Lancet Diabetes Endocrinol*. 2023 Sep;11(9):644-656. https://doi.org/10.1016/S2213-8587(23)00171-7.
21 Palileo-Villanueva LM, Dans AL. Composite endpoints. *J Clin Epidemiol*. 2020 Dec;128:157-158. https://doi.org/10.1016/j.jclinepi.2020.07.017.
22 Multiple Endpoints in Clinical Trials Guidance for Industry. Multiple Endpoints in Clinical Trials. Guidance for Industry. U.S. Department of Health and Human Services. Food and Drug Administration. Center for Drug Evaluation and Research (CDER). Center for Biologics Evaluation and Research (CBER). October 2022. *Biostatistics*. (PDF accessed 25 March 2024).
23 U.S. Department of Health and Human Services, Food and Drug Administration (FDA). Guidance for Industry Patient Reported Outcome Measures: Use in Medical Product Development to Support Labelling Claims. 2009. (PDF accessed 8 May 2023).
24 Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Annals of Internal Medicine*. 1993;15;118(8):622-629. https://doi.org/10.7326/0003-4819-118-8-199304150-00009.
25 Bergner M, Bobbitt RA, Carter WB, et al. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care*. 1981;19(8):787-805.
26 Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*. 1992;30(6):473-483. http://www.jstor.org/stable/3765916.
27 EuroQol–a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy*. 1990;16(3):199-208. https://doi.org/10.1016/0168-8510(90)90421-9.
28 Zhang AD, Ross JS. Biomarkers as Surrogate Endpoints: Ongoing Opportunities for Validation. *Journal of Law, Medicine & Ethics*. 2019;47(3):393-395. https://doi.org/10.1177/1073110519876170.
29 MF Drummond, MJ Sculpher, K Claxton, GL Stoddart, GW Torrance. Methods for the economic evaluation of health care programmes. Fourth Edition. *Oxford: Oxford Medical Publications*, 2015. (PDF accessed 27 February 2024).
30 K. Mccabe, E. Kirwin, J. Round, K. Bond; A Conceptual Framework for Life-Cycle Health Technology Assessment; *Value in Health* 25(7); 2022 A Conceptual Framework for Life-Cycle Health Technology Assessment. https://doi.org/10.1016/j.jval.2021.11.1373.

31 Dean R, Jensen I, Cyr P, Miller B, Maru B, Sproule DM, Feltner DE, Wiesner T, Malone DC, Bischof M, Toro W, Dabbous O. An updated cost-utility model for onasemnogene abeparvovec (Zolgensma®) in spinal muscular atrophy type 1 patients and comparison with evaluation by the Institute for Clinical and Effectiveness Review (ICER). *J Mark Access Health Policy*. 2021 Feb 28;9(1):1889841. https://doi.org/10.1080/20016689.2021.1889841.

32 Modified from: Schneeweiss S, Patorno E. Conducting real-world evidence studies on the clinical outcomes of diabetes treatments. *Endocrine Reviews*. 2021;42(5):658-690. https://doi.org/10.1210/endrev/bnab007.

33 Schneeweiss S, Patorno E. Conducting real-world evidence studies on the clinical outcomes of diabetes treatments. *Endocrine Reviews*. 2021;42(5):658-690. https://doi.org/10.1210/endrev/bnab007.

34 Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560. https://doi:10.1097/00001648-200009000-00011.

35 Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000:561-570. https://doi:10.1097/00001648-200009000-00012.

36 Weuve J, Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology*. 2012;23(1):119-128. https://doi:10.1097/EDE.0b013e318230e861.

37 Schneeweiss S, Patorno E. Conducting real-world evidence studies on the clinical outcomes of diabetes treatments. *Endocrine Reviews*. 2021;42(5):658-690. https://doi.org/10.1210/endrev/bnab007.

38 Azoulay L, Suissa S. Sulfonylureas and the risks of cardiovascular events and death: a methodological meta-regression analysis of the observational studies. *Diabetes Care*. 2017;40(5):706-714. https://doi.org/10.2337/dc16-1943.

39 Suissa S. Immortal time bias in pharmacoepidemiology. *American Journal of Epidemiology*. 2008;167(4):492-499.https://doi.org/10.1093/aje/kwm324.

40 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Current Epidemiology Reports*. 2015;2(4): 221-228. https://doi.org/10.1007/s40471-015-0053-5.

41 Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *American Journal of Epidemiology*. 2003;158(9):915-920. https://doi.org/10.1093/aje/kwg231.

42 Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care*. 2012;35(12):2665-2673. https://doi.org/10.2337/dc12-0788.

43 Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiology and Drug Safety*. 2017;26(4):459-468. https://doi.org/10.1002/pds.4107.

44 Khosrow-Khavar F, Bouganim N, Filion KB, Suissa S, Azoulay L. Cardiotoxicity of use of sequential aromatase inhibitors in women with breast cancer. *American Journal of Epidemiology*. 2020;189(10):1086-1095. https://doi.org/10.1093/aje/kwaa065.

45 Khosrow-Khavar F, Filion KB, Al-Qurashi S, Torabi N, Bouganim N, Suissa S, Azoulay L. Cardiotoxicity of aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer: a systematic review and meta-analysis of randomized controlled trials. *Annals of Oncology*. 2017;28(3):487-496. https://doi.org/10.1093/annonc/mdw673.

46 Khosrow-Khavar F, Filion KB, Bouganim N, Suissa S, Azoulay L. Aromatase inhibitors and the risk of cardiovascular outcomes in women with breast cancer: a population-based cohort study. *Circulation*. 2020;141(7):549-559. https://doi.org/10.1161/CIRCULATIONAHA.119.044750.

47 Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000:561-570. https://doi:10.1097/00001648-200009000-00012.

48 Weuve J, Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology*. 2012;23(1):119-128. https://doi:10.1097/EDE.0b013e318230e861.

49 Gruber S, Lee H, Phillips R, Ho M, van der Laan M. Developing a Targeted Learning-Based Statistical Analysis Plan. *Statistics in Biopharmaceutical Research*. Published online October 3, 2022:1-8. https://doi.org/10.1080/19466315.2022.2116104.

50 Gruber S, Phillips RV, Lee H, Ho M, Concato J, van der Laan MJ. Targeted learning: Toward a future informed by real-world evidence. *Stat. Biopharm*. Res. 2022; https://doi.org/10.1080/19466315.2022.2116104.

51 Concato J, Corrigan-Curay J. Real-world evidence—Where are we now? *New Engl J Med* 2022;1680-1682. DOI: 10.1056/NEJMp2200089.

52 Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. *Clinical Pharmacology & Therapeutics*. 2020;107(4):817-826. https://doi.org/10.1002/cpt.1633.

53 PRECIS-2. (Website accessed October 2022).

54 The United States Food and Drug Administration (FDA). Determining the extent of Safety Data Collection in Late-Stage Premarket and Post-approval Clinical Investigations. (Website accessed 2 May 2023).

55 European Medicines Agency, Reflection paper on establishing efficacy based on single-arm trials submitted as pivotal evidence in a marketing authorisation, Considerations on evidence from single-arm trials. 17 April 2023. (PDF accessed 27 February 2024).

56 Codes of Federal Regulations 21CFR 314.126. Adequate and well-controlled studies. (Website accessed 9 May 2023).

57 European Medicines Agency (EMA). ICH E10 Choice of control group in clinical trials - Scientific Guideline. (PDF accessed 9 May 2023).

58 The United States Food and Drug Administration (FDA). Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. Guidance for Industry.Draft guidance. (PDF accessed 9 May 2023).

59 European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). ENCePP Guide on Methodological Standards in Pharmacoepidemiology. (Website accessed October 2022).

60 Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008;20:106-107.

61 Statology. What is Referral Bias? (Website accessed October 2022).

62 Eysenbach G, Wyatt J. Using the Internet for surveys and health research. *Journal of Medical Internet Research*. 2002;22;4(2):e862. https://doi:10.2196/jmir.4.2.e13.

63 Neyman J.. Statistics; servant of all sciences. *Science*. 1955;122(3166):401–406. https://doi:10.1126/science.122.3166.401.

64 Chubak J, Boudreau DM, Wirtz HS, McKnight B, Weiss NS. Threats to validity of nonrandomized studies of postdiagnosis exposures on cancer recurrence and survival. *Journal of the National Cancer Institute*. 2013;2;105(19):1456-1462. https://doi.org/10.1093/jnci/djt211.

65 Horwitz RI, Feinstein AR. Alternative analytic methods for case-control studies of estrogens and endometrial cancer. *New England Journal of Medicine*. 1978;16;299(20):1089-1094. https://doi/10.1056/NEJM197811162992001.

66 Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008;20:106-107.

67 Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiology and Drug Safety*. 2007;16(3):241-249. https://doi.org/10.1002/pds.1357.

68 Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *American Journal of Epidemiology*. 2008;1;168(3):329-335. https://doi.org/10.1093/aje/kwn135.

69 Pagano M, Gauvreau K. *Principles of Biostatistics*. 2nd edition. Pacific Grove, CA: Duxbury; 2000.

70 JP Klein, JD Rizzo, M-J Zhang and Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. *Bone Marrow Transplantation*. 2001;28:909-915. https://doi.org/10.1038/sj.bmt.1703260.

71 Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC medical research methodology*. 2008 Dec;8:1-5. https://doi.org/10.1186/1471-2288-8-70.

72 Cohen JF, Korevaar DA, Altman DG et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;14;6(11):e012799. http://dx.doi.org/10.1136/bmjopen-2016-012799.

73 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94:496-509. https://doi.org/10.2307/2670170.

74 Aalen O, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand Stat Theory Appl*. 1978; 5(3): 141- 150.

75 Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*. 1999;30;18(6):695-706. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O.

76 Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics*. 1988;16:1141-1154.

77 Cox DR. Regression model and life tables. Journal of the Royal Statistical Society. Series B (Methodological). 1972;34:187-200.

78 JP Klein, JD Rizzo, M-J Zhang and Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part 2: Regression modeling. *Bone Marrow Transplantation*. 2001;28:1001-1011. https://doi.org/10.1038/sj.bmt.1703271.

79 Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*. 2013;15;177(4):292-298. https://doi.org/10.1093/aje/kws412.

80 Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ*. 2017 Oct 16;359. DOI: 10.1136/bmj.j4587.

81 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;1;70(1):41-55. https://doi.org/10.1093/biomet/70.1.41.

82 Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*. 1984;1;79(387):516-524. https://doi:10.1080/01621459.1984.10478078.

83 Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*. 1985;39:33-38. https://doi:10.1080/00031305.1985.10479383.

84 Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of Epidemiology & Community Health*. 2006;1;60(7):578-586. http://dx.doi.org/10.1136/jech.2004.029496.

85 Velentgas P, Dreyer N, Noujah P, Smith SR, and Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Agency for Health Care Research and Quality (U.S.). Government Printing Office, 2013/02/21.

86 Greifer N, Stuart EA. Matching methods for confounder adjustment: an addition to the epidemiologist's toolbox. *Epidemiologic Reviews*. 2022;43(1):118-129. https://doi.org/10.1093/epirev/mxab003.

87 Gruber S, Phillips RV, Lee H, Ho M, Concato J, van der Laan MJ. Targeted learning: Toward a future informed by real-world evidence. *Stat. Biopharm. Res.* 2022; https://doi.org/10.1080/19466315.2022.2116104.

88 Dorie V, Hill J, Shalit U, Scott M, and Cervone D. Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition. July 2017; *Statistical Science* 34(1). DOI: 10.1214/18-STS667.

89 Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol* 2011 Apr;173(7):761-7. [PMCID: PMC3070495]. DOI: 10.1093/aje/kwq439.

90 European Medicines Agency (EMA).ICH E9 (R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. (PDF accessed 2 May 2023).

91 Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and Drug Safety*. 2006;15(5):291-303. https://doi.org/10.1002/pds.1200.

92 Velentgas P, Dreyer N, Noujah P, Smith SR, and Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Agency for Health Care Research and Quality (U.S.) Government Printing Office; 2013.

93 Lu S, Ding P. «Flexible sensitivity analysis for causal inference in observational studies subject to unmeasured confounding. 2023; arXiv:2305.17643. https://doi.org/10.48550/arXiv.2305.17643.

94 Zhang X, Stamey JD, Mathur MB. Assessing the impact of unmeasured confounders for credible and reliable real-world evidence. *Pharmacoepidemiol Drug Saf*. 2020 Oct;29(10):1219-1227. https://doi.org/10.1002/pds.5117.

95 Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*. 2012;4;367(14):1355-1360. https://doi:10.1056/NEJMsr1203730.

96 Little RJ, Rubin DB. Statistical analysis with missing data. John Wiley & Sons; 2019.

97 Lee KJ, Tilling K, Cornish RP, Little RJ, Bell ML, Goetghebeur E, et al. for the STRATOS initiative. Framework for the treatment And reporting of missing data in observational studies: the TARMOS framework. *Journal of Clinical Epidemiology*. 2021;1;134:79-88. https://doi.org/10.1016/j.jclinepi.2021.01.008.

98 Website access 4 June 2023.
99 Orsini LS, Berger M, Crown W, Daniel G, Eichler HG, Goettsch W, et al. Improving transparency to build trust in real-world secondary data. Studies for hypothesis testing—why, what, and how: recommendations and a road map from the real-world evidence transparencyinitiative. *Value in Health*. 2020;23(9):1128-1136. https://doi.org/10.1016/j.jval.2020.04.002.
100 Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Value in Health*. 2017;20(8):1003-1008. https://doi.org/10.1016/j.jval.2017.08.3019.
101 Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Value in Health*. 2017;20(8):1003-1008. https://doi.org/10.1016/j.jval.2017.08.3019.
102 Peat G, Riley RD, Croft P, Morley KI, Kyzas PA, Moons KG, et al. Improving the transparency of prognosis research: the role of reporting, data sharing, registration, and protocols. *PLoS Medicine*. 2014;11(7):e1001671. https://doi.org/10.1371/journal.pmed.1001671.
103 Sox HC, Helfand M, Grimshaw J, Dickersin K, PLoS Medicine Editors, Tovey D, et al. Comparative effectiveness research: challenges for medical journals. *PLoS Medicine*. 2010;7(4):e1000269. https://doi.org/10.1371/journal.pmed.1000269.
104 Williams RJ, Tse T, Harlan WR, Zarin DA. Registration of observational studies: is it time? *Canadian Medical Association Journal*. 2010;182(15):1638-1642. https://doi.org/10.1503/cmaj.092252.
105 Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I, et al. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Value in Health*. 2017;1;20(8):1009-1022. https://doi.org/10.1016/j.jval.2017.08.3018.
106 Mullard A. Half of cancer studies fail high-profile replication test. *Nature*. 2021;600:368-369. https://doi.org/10.1038/d41586-021-03691-0.
107 Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*. 2015;28;349(6251):aac4716. https://:doi:10.1126/science.aac4716.
108 Camerer CF, Dreber A, Forsell E, Ho TH, Huber J, Johannesson M, et al. Evaluating replicability of laboratory experiments in economics. *Science*. 2016;25;351(6280):1433-1436. https://doi:10.1126/science.aaf0918.
109 Ebrahim S, Sohani ZN, Montoya L, Agarwal A, Thorlund K, Mills EJ, et al. Reanalyses of randomized clinical trial data. *JAMA*. 2014;10;312(10):1024-1032. https://doi:10.1001/jama.2014.9646.
110 Wang SV, Sreedhara SK, Schneeweiss S. Reproducibility of real-world evidence studies using clinical practice data to inform regulatory and coverage decisions. *Nature Communications*. 2022;31;13(1):5126. https://doi.org/10.1038/s41467-022-32310-3.
111 Wang SV, Pottegård A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, et al. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force. *Value in Health*. 2022;1;25(10):1663-1672. https://doi.org/10.1016/j.jval.2022.09.001.
112 Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England Journal of Medicine*. 2000;342(25):1887-1892. https://doi:10.1056/NEJM200006223422507.
113 Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. *Clinical Pharmacology & Therapeutics*. 2020;107(4):817-826. https://doi.org/10.1002/cpt.1633.
114 Wang SV, Schneeweiss S; RCT-DUPLICATE Initiative; Franklin JM, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023 Apr 25;329(16):1376-1385. doi: 10.1001/jama.2023.4221. Erratum in: *JAMA*. 2024 Apr 9;331(14):1236.
115 RCT DUPLICATE. (Website accessed 2 June 2023).
116 Observational Patient Evidence for Regulatory Approval and uNderstanding Disease (OPERAND). (Website accessed 2 June 2023).
117 Center of Excellence in Regulatory Science and Innovation (CERCI). (Website accessed 2 June 2023).
118 Wang SV, Schneeweiss S; RCT-DUPLICATE Initiative; Franklin JM, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023 Apr 25;329(16):1376-1385. doi: 10.1001/jama.2023.4221. PMID: 37097356; PMCID: PMC10130954.
119 Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023;329(16):1376–1385. doi:10.1001/jama.2023.4221.
120 Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, Strobe Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Annals of Internal Medicine*. 2007;147(8):573-577. https://doi.org/10.7326/0003-4819-147-8-200710160-00010.
121 Langan SM, Schmidt SA, Wing K, Ehrenstein V, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018 Nov 14;363:k3532. https://doi.org/10.1136/bmj.k3532.
122 Benchimol EI, Smeeth L, Guttmann A, Harron K, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015 Oct 6;12(10):e1001885. https://doi.org/10.1371/journal.pmed.1001885.
123 British Medical Journal (BMJ). BMJ Best Practice. What is Grade? (Website accessed 15 October 2022).
124 Schünemann HJ, Cuello C, Akl EA, Mustafa RA, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol*. 2019 Jul;111:105-114. https://doi.org/10.1016/j.jclinepi.2018.01.012. Epub 2018 Feb 9.
125 Cuello-Garcia CA, Santesso N, Morgan RL, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. *Journal of Clinical Epidemiology*. 2022 Feb;142:200-208. https://doi.org/10.1016/j.jclinepi.2021.11.026.
126 Balshem H, Helfand M, Schünemann HJ, Oxman AD, GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011 Apr;64(4):401-6. https://doi.org/10.1016/j.jclinepi.2010.07.015. Epub 2011 Jan 5.
127 Schünemann HJ, Cuello C, Akl EA, Mustafa RA, et al. GRADE Working Group. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol*. 2019 Jul;111:105-114. https://doi.org/10.1016/j.jclinepi.2018.01.012. Epub 2018 Feb 9.

128 Cuello-Garcia CA, Santesso N, Morgan RL, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. *Journal of Clinical Epidemiology*. 2022 Feb;142:200-208. https://doi.org/10.1016/j.jclinepi.2021.11.026.
129 Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ*. 2016 Jun 30;353:i2089. https://doi.org/10.1136/bmj.i2089.
130 Sterne JA, Hernán MA, Reeves BC, Savović J, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016 Oct 12;355:i4919. https://doi.org/10.1136/bmj.i4919.
131 Allen A, Patrick H, Ruof J, Buchberger B, et al. Development and Pilot Test of the Registry Evaluation and Quality Standards Tool: An Information Technology-Based Tool to Support and Review Registries. *Value in Health*. 2022 Aug;25(8):1390-1398. https://doi.org/10.1016/j.jval.2021.12.018. Epub 2022 Mar 8.

CHAPTER 4.
ETHICS AND GOVERNANCE

Two situations point to some of the limitations of the current gold standard of RCTs. First, the reality of modern medicine development is that we need to start considering different data sources in addition to RCT data to support our evidence generation for medicine development. RCTs may not be practical in certain populations, such as patients suffering from rare diseases, and in scenarios where there is a promising treatment with high unmet need. Second, the requirements for documenting efficacy tend to limit the participants in studies to a group that excludes vast segments of the population, and more specifically, excludes under-served populations such as patients from ethnic minorities, those from older age groups, and those with comorbidities (some of these are also harder to recruit to trials).

The Introduction through to Chapter 3 of the report show strong evidence of a compelling scientific case for an extended utilisation of RWD, including data addressing specifically efficacy/effectiveness, which is no longer an exclusive domain of RCTs, and this includes change at the regulatory level. Indeed, this has already begun, with some jurisdictions moving to include RWE in key statutes around the regulation of medicine development.

While RCTs are likely to remain an essential part of the evidence base, there is scientific argument for increasing the use of RWE to augment the available clinical evidence. However, that scientific argument poses questions in law and ethics. This chapter addresses questions about the normative implications of the change to incorporating more RWE, namely:

1. Is it ethical to continue assessment of medicinal products without integrating other forms of evidence when available?
2. What ethical and legal issues need to be considered when using more RWD?

To understand these questions, it is necessary to discuss some underpinning ethics concepts, particularly the nature of duty (and who owes duties to whom), and the nature of autonomy and solidarity. There are also a number of ethical and legal questions to address, particularly around the protection of personal data and the secondary use of already gathered data.

A number of fundamental questions about data sharing norms must also be considered, particularly the nature of privacy rights, and how far informed consent is required for the re-use of personal data in settings different from where it was initially gathered. There are also fundamental questions about how data and evidence about individuals in real-world settings are constructed, so that the observed-self retains all its legal rights, remains autonomous both as an individual and in community with the general population. Within these broader questions, all voices that wish to contribute to the discussion must be considered.

The aim of the chapter is not necessarily to produce definitive answers to these issues. Rather, this chapter contributes to the framing of the discussion about how to respond to the introduction and the first three chapters of the report.

4.1. Ethical arguments for incorporating real-world evidence

The broader use of RWE to evaluate efficacy as well as safety is justified not only by a need for stronger and broader evidence and to include neglected groups in the evidence base, but also by concerns that evidence from RCTs often does not translate into real-world use. In other words, the evidence on efficacy from RCTs may not translate into evidence of effectiveness in clinical care. This is because the actual patient population is often not well represented by participants in RCTs, who are often younger and healthier than many patient groups in daily practice. Clinical trials also tend to under-report harm, given the narrow eligibility that limits exposure to certain groups until safety is proven, further weakening the evidence base for real-world clinical care.[1]

This phenomenon is known as the efficacy-effectiveness gap; evidence shows that the efficacy-effectiveness gap worsens disease response and survival outcomes and increases toxicity in the clinical setting.[2] Patients treated in everyday practice tend to be older and more frail, to have poorer function and performance status, and to have more comorbidities and less social support than those selected to participate in clinical trials.

Informed decision making with patients mostly relies on evidence from clinical trials that describe the likely benefits and safety. Thus, generalisability to typical patient populations treated in daily practice is often limited. Kennedy-Martin and colleagues explored the generalisability of RCTs in cardiology, mental health, and oncology by assessing studies comparing participants in such trials with those in everyday clinical practice.[3] Patients treated in everyday clinical practice tended to be older, were more often women, and had more comorbidities; 71% of studies concluded explicitly that RCTs were not broadly representative of real-world patients; in particular, pregnant and lactating women are a very large population that is often entirely unrepresented in clinical trials.

Furthermore, patients enrolled in trials were more likely treated according to guidelines and were subject to additional procedures required to monitor patients during trials. Strict selection criteria for RCTs meant that participants were at a much lower risk of adverse events compared with patients treated in clinical practice.

4.2. Ethical issues in using real-world data

Before considering privacy and data protection concerns regarding the use of RWD, we should note that relying more on RWD also carries its own potential disadvantages. For example, RWE faces challenges, due to the potential for unmeasured confounding, limitations on clinical practice such as use outside that intended (off-label use), and potential lack of adherence with treatment.

Furthermore, while RCTs suffer from the disadvantages of non-representativeness, neglect of underserved groups, and the efficacy-effectiveness gap, RCTs do have the advantage of being designed to control for confounders and other biasing factors; indeed, this is one of the reasons why randomisation and control are seen as being so important. If RWD is to be used more, the potential for biases, confounders and other weaknesses in the RWE derived from RWD must be acknowledged in decision making.

Similarly, as mentioned in Chapter 3, reproducibility may be more of a concern for RWD than for RCTs.

While RWD constitutes a resource with great potential, that potential can only be realised if the RWE derived from those data is reliable, representative and robust. If unreliable RWD and RWE were used for decision making, the problems with RCTs would simply be replaced with a new set of problems,

resulting in an equally flawed evidence base. As stated in Chapter 1, it is likely that an evaluation of the methods used to generate the RWE, along with the reliability and relevance of the RWD involved, will play a central role. In any case, this must be borne in mind as a potential ethical issue.

In addition to privacy and data protection, perhaps the most important ethical issue concerning use of RWD is informed consent. In many cases, patient data is routinely used for service evaluation and audit without explicit consent being sought. For example, some HCPs in the UK simply display posters informing patients about this. If RWD is to be used more, then routine data linkage with patient records for research in addition to evaluation and audit may be a next step.

As mentioned in Chapter 2, such linkage raises some issues regarding confidentiality and privacy. However, given that such data is often used without explicit consent for other purposes, it might be argued that seeking active consent for research use is disproportionate given the potential benefits of using such data. In pragmatic clinical trials and comparative effectiveness trials, it is already accepted that consent may not be necessary where multiple standard-of-care interventions are being compared.[4]

In any case, if RWD is to be used in a way that is truly representative of populations and underserved groups, enabling people to opt their data out of RWE generation efforts may be counterproductive to the goal of generating decision-grade evidence. However, any such change in paradigm should not be accomplished by diktat; societal discussion should precede any such legislative change. Depending on the jurisdiction, some model of opt-out or dynamic consent, where patients can closely control how their data are used over time, may be feasible.

Informed consent is also an issue in pragmatic trials where RWD sources are used. As stated in Chapter 3, these are often integrated with routine practice, in a context where explicit consent (clear, direct consent between two parties) is not necessary provided there is sufficient ethical oversight. This is justified because randomisation is not occurring at the individual patient level. Depending on the specific context, simplified consent (consent with the minimal information necessary, usually for procedures or practice with minimum risks) may still be appropriate.

One situation in which consent is less of an issue is the use of wearable biosensors such as the Apple Watch (see Chapter 2). Users of such devices can opt in to share their data with manufacturers and associated research teams, with the Apple Heart Study gathering RWD from over 400 000 participants.[5] Encouraging people to opt in to such studies will increase awareness of the importance of RWD and may help move the societal paradigm towards greater acceptance of routine data sharing and linkage. As mentioned in Chapter 2, the use of social media data to track adverse events is also a promising use of RWD. In that context, too, users have directly shared data publicly about adverse events in ways that are beneficial to research.

The issue of consent to data use if closely linked to the issues of privacy and data protection, which are explored in the next section.

4.3. Privacy and data protection

The purpose of data protection legislation is to protect the fundamental rights and interests of citizens in relation to the processing of personal data that relate to them. However, this can be satisfied in many situations where sensitive personal data about individuals are processed, for example, in relation to banking details or in other commercial transactions that place citizens in vulnerable situations in relation to their personal data, through the safeguard of a clickwrap or click-to-sign consent. It is an informed consent, but it is un-negotiated, and often largely unread, including lengthy terms that

seem to offer no realistic safeguard for the individual to protect personal data, or for that personal data to be properly protected in the transaction.

On the other hand, highly regulated areas such as medical research, with multiple safeguards and independent scrutiny are made almost impossible to negotiate. RWD is in danger of being so restricted by data protection laws and regulations that it becomes impossible to work with, whereas in practice it is an area where the interests of individual citizens are robustly protected, more so than in many commercial situations imposed on consumers, and where the outcomes that the RWD research pursues are clearly in the public interest and in the interests of protecting human dignity.

4.3.1. Data protection landscape, using the EU legislation as a case study

From its common international roots in the late 1970s,[6] data protection law has shared a common language and basic shape.[7] The underpinning idea is that the individual citizen has human rights, particularly privacy rights in relation to the processing of their personal data. These are expressed primarily in duties imposed on those who process personal data (or who have obligations flowing from someone with such duties), and actionable rights on the part of the individuals to whom the data relate (data subjects). Persons with duties can be both natural persons (human, physical people) and legal persons (e.g. companies). Individuals to whom those duties are owed, interestingly, tend to be individuals and not groups of individuals.

In the following explanation of the rights and duties, we are using the European Union's General Data Protection Regulation (EU GDPR 2016/679) as an example.

The duties owed by those who process personal data (particularly by those who determine how data will be processed and for what purposes it will be processed) are captured in data protection principles: to process the data fairly, lawfully and in a transparent manner (i.e. the processes will be transparent); to process the data for specified purposes and not thereafter for purposes that are incompatible with those initial stated purposes; to minimise the data that is processed (i.e. only to collect and process data necessary for the purpose of the processing); to keep data only for so long as is necessary for the purposes of the processing; to keep the data secure; and to act with integrity towards the data.

Lawful processing is prescribed to include, although not exclusively, two fundamental elements:

1. processing must be on the legal bases for the processing of personal data; and
2. data subjects must be given information about the identity and contact details of the data controller and the purpose and nature of the processing of the personal data.

The GDPR includes several further obligations, for example, the duty of data protection by design, ensuring that any activities including the processing of personal data consider the implications of data protection expectations from the outset, and administrative structures for the enforcement of (considerable) sanctions in the case of breach.

The Article 35 duty of data protection by design seeks to ensure that data protection considerations are embedded in any enterprise processing personal data from the outset. For RWD this not only invites the question, are the data being processed personal data, and therefore within the ambit of the law? It also asks, what are the data science mechanisms that are available, or need to be developed, that ensure the protection of personal data?

Data protection by design arguably invites a pro-active approach to addressing the protection of data subjects rather than ensuring a minimal compliance. For example, how far can data science tools show definitively who has had access to personal data as it is processed?

4.3.2. Data protection and processing real-world data

Here, we consider the major unresolved conceptual and technical issues for the use of RWD.

Strong data protection law should not be seen as an obstacle or barrier to the effective processing of personal data, and therefore where there are unresolved technical issues they must be resolved. There is a strong argument that the processing of RWD only works where data subjects have trust and confidence in the institutions and individuals who process data that relate to them, and therefore a strong and well-functioning personal data protection regime is essential to the acceptance and operation of RWD processing.

However, to be effective and to foster trust and confidence, the data protection regime must equally be coherent, appropriate and transparent. It must be coherent across the sector; trials and biomedical research must operate at an international level, and there needs to be a very strong argument for the regulatory frameworks to operate seamlessly across jurisdictions. This requires political will to discuss and understand different perspectives and concerns to ensure that the range of safeguards put in place internationally reflect the concerns of individuals and their communities.

The measures must be appropriate in that they must reflect the balance of interests at stake in the sector. These interests are complex and can seem conflicting: citizens at the same time hold aspirations and concerns about the development of new therapies to treat, cure and prevent illness, and about their privacy and use of their personal data in different contexts; whilst industries seeking to process RWD with the commendable aim of therapy development can appeal to an altruism underpinning their motives, they must also acknowledge that their work is also designed to make profit in a commercial environment, and that personal data can easily become a commodity.

Regulatory measures must acknowledge and balance these tensions, and again, there must be a political will to create that balance. It goes without saying that the measures taken must be effective, but considering this requirement, and reflecting on the other two elements, competing expectations between all the parties must be properly managed. For example, data subjects cannot expect cutting-edge pharmaceutical product development but also completely opt out of allowing the use of data that relate to them in the development of such products; companies cannot expect unfettered access to personal data on the basis of the public interest or a simple consent, and must respect the need for equitable access to products.

There is a strong case that to be successful and incentivise the public, RWD relies on an altruistic society that must be realised through its regulatory and governance structures. However, these are part of a social and political debate that is necessary to underpin the interpretation and operation of data protection law, and this debate is yet to be had in the broader community.

4.3.3. Legal basis

Like all gathering of personal data, RWD are gathered with a legal basis for processing. The nature of RWD is that it can be a collection of already gathered data that are repurposed (further processed) for a new situation. They are gathered from many sources and combine different data sets.

Unfortunately, data protection law is conceptually focused on what might be described as single-purpose processing. Personal data, in classical data protection thinking, are gathered for a purpose or purposes that are discerned at the outset of the project, and whilst the legislation allows for further processing for novel purposes that were not imagined at the outset, it is not easily negotiated, as will be seen. This is the opposite of RWD processing, which is concerned about previously unimagined and novel deployment of data.

Much medical data is gathered either on the basis of informed consent or on the basis of an implied consent through the general contract between a HCP and a patient. On the former point, research ethics committees (RECs) and the general operation of patient rights and bioethics has set up the expectation that informed consent is the expected legal basis for medical interactions, and this has reached into personal data processing as an expression of autonomy.

In many jurisdictions, personal data are gathered as part of the protection of both patients and medical professionals, on the basis of the statutory duty to create a medical record for each patient. In this case, how far the duty is drawn to the attention of the patient in the creation of the relationship at the outset is one issue. The more problematic issue is that the same patient rights statutes that create this duty also create duties of confidentiality for the processing of the medical data and record, that limit the transfer of data to necessary transfers within the clinical context. To stretch this to the research context is difficult; we return to this issue below.

The original gathering of personal data, i.e. data that relate to an identified or identifiable individual, is on the basis of a particular purpose. Therefore, the first question relating to the processing of RWD is: does the original legal basis for processing cover a new, unforeseen purpose for further processing? This is complicated by the tendency for modern data protection to see informed consent as narrow or specific. The opportunities for broad consent are made within, for example, the GDPR, but they are not explained clearly in the heart of the legislation, and the individual member states of the EU have shown that there are considerable differences in both the technical and conceptual willingness to explore broad consent for research fully. As indicated, other relevant RWD will be gathered on the basis of the statutory requirements of patient rights and medical practice, or perhaps on the basis of necessity, for example in the emergency room. Therefore, the answer to our first question could well be that the original legal basis does not cover the proposed new processing.

4.3.4. Compatible processing

As indicated above, all is not lost at this point. The GDPR indicates that personal data should be gathered for an identifiable purpose or purposes and not further processed for incompatible purposes. Therefore, processing for purposes that are compatible with the purpose of the original gathering and processing of the data are permitted. In addition, the GDPR goes further to indicate that further processing for research purposes are compatible with the original purpose.

In the case of the GDPR, this is very positive for RWD processing. However, it is not without difficulties. Research under the GDPR includes applied research, so the activities of pharmaceutical industries, for example, would be included. However, where the data have been gathered under the statutory duty to create a patient record, with the requirements that such data be treated confidentially within the clinical setting, we will face the argument that using these data in RWD settings is incompatible with the original purpose. This would be because the research processing is incompatible with the original purpose. It is a question of the hierarchy of the laws in place.

The same issue arises in relation to informed consent situations. Where an informed consent has explicitly excluded the proposed further processing, can this new processing be undertaken as compatible? Arguably, it is explicitly incompatible, even in the face of the statutory presumption to

the contrary. These issues must be resolved. Of course, the easiest way to resolve the issue is to include the possibility of future processing for RWD research settings in the legal basis upon which new data are gathered from now. However, RWD contains historical data, and the prospective solution is therefore not sufficient.

4.3.5. Information provision

Separately to the requirement for a legal basis for processing, those who process personal data must inform the data subjects of their identity, contact details and the purpose for and nature of the processing they propose. This is not a requirement for informed consent in all cases. It acknowledges that the data subject has rights that they can only engage when they are aware that processing is taking place. It allows, in certain circumstances, for data subjects to opt-out or modify their participation in certain processing, and is therefore a necessary part of the process.

Under the GDPR, a distinction is made concerning the information that must be provided by a researcher to a data subject, e.g. to inform the participant about the data processing activities, between direct and indirect gathering of personal data: when the researcher gathers data directly from a data subject, information must be provided to the subject (GDPR Article 13), whereas when the data are gained indirectly, i.e. from another source, then the expectation is that the information must be provided unless it is impossible or requires a disproportionate effort (GDPR Article 14). It should be noted that where the data are gained indirectly, this is likely to be from a data controller who has gathered the data directly from the data subject.

As mentioned in Chapter 2, genome databases and biobanks raise important concerns about data ownership and control, with researchers favouring open use of such data and samples. But companies and governments often wish to control access to such data for commercial reasons. Data ownership is complicated by the fact that genetic data concern not only the data subject, but also the data subject's family members. Great care must be taken when processing genomic and genetic information. Of particular interest are genetic relatives of donors to biobanks, whose data will be included indirectly but without a direct gathering data controller.

This, arguably, does not cause a difficulty, except for compatible processing in RWD scenarios. Where the data controller has gathered data originally from the data subject and then seeks to process those data in a RWD secondary processing, the controller must inform the data subject of this new, compatible processing. The same applies where the data are gathered from another data controller. In that scenario, the recipient data controller can rely on the caveat for indirect processing (impossibility or disproportionate effort). However, the original data controller must inform the data subject of the transfer, unless it was explained in the original information provided at the gathering of the data.

Key here is what is an acceptable way of informing the individual data subject of the compatible or otherwise secondary processing of personal data that relate to them. Where this is on the basis of direct informing, the costs and possibility of doing so in a RWD scenario are likely to make the enterprise too costly.

4.3.6. De-identifying the data

Data protection law only operates on personal data, meaning data that identify an individual natural person or that are capable of doing so when linked to other data, something one might term mosaicking. The easiest example to comprehend is pseudonymised data. Personal data have certain identifiers, for example a name or address, which are replaced with a code. 'Pseudonymisation' of

data means replacing any information which could be used to identify an individual with a pseudonym, or, in other words, a value which does not allow the individual to be directly identified.

The effect pseudonymisation is that the remaining dataset (the coded data) does not of itself disclose the identity of the individual to whom the data relate. However, the code is kept elsewhere and when it is reunited with the rest of the dataset, the whole dataset can re-identifying the individuals. Data protection law sees pseudonymised data — both the code and the coded data in the example — as all being personal data; all the pseudonymised data are capable of being combined to identify individuals.

The question is one of the likelihood of the reconnection of the data. Some jurisdictions have taken a view that, once de-identified in this way, even when identifiable data are still available elsewhere (for example if a sample of data is copied from a biobank and given to a researcher in a de-identified form, with the data still existing in an identifiable form in the biobank), the break will be sufficiently to render the data as de-identified and no longer personal in the hands of the researcher. Other jurisdictions take a harder line, whereby the very possibility of reconnecting the de-identified data with the identifying data will maintain the personal quality of the data in the hands of the researcher who has received de-identified data.

This is another area where policy must be considered and then harmonised. The GDPR, using the idea of reasonableness in assessing the possibility of re-identification shows pragmatism in the letter of the law, but requires harmonisation in the interpretation of the idea to ensure consistency.

An interesting aspect in the use of de-identified data is where it is linked to federated data projects. Imagine a research project where data are de-identified by several data controllers and those data are then passed to researchers. In many jurisdictions this would remove the data from the scope of the data protection law. However, the data remain identifiable in the hands of the original data controllers. The researchers could run into a question about their dataset and send a question to the data controllers from whom they received the data. As the data are still identifiable by the original data controller, the question is whether the data controllers can respond to the researcher's question with a de-identified response. In this sort of case, at what point does the de-identification become an arms-length pseudonymisation? By whom and when will this be questioned or regulated?

The current law has created a strange situation where workarounds are tried against the backdrop of differing approaches by data protection officers, data stewards, RECs and institutional review boards (IRBs). Too often, however, they can get very little guidance from the regulators before interventions for breaches are made.

4.3.7. Research or safety evaluation

In clinical trials and medicine production, it is very interesting that much of what has been discussed above in relation to research does not apply to the conduct of evaluations for safety of medicines on the market. In this case, public safety conceptually trumps individual privacy or autonomy claims. While this fits with the legal basis, as processing for the public interest and for statutory duty is well established, it is not easily reconciled with the information provision.

However, whereas most legislation that regulates situations where personal data are processed defer to the GDPR to govern the processing of personal data, for example, the Clinical Trials Regulation in the EU, it is possible for safety governance to overrule the general data protection legislation. This makes for an interesting anomaly in RWD processing: that processing for research must be GDPR compliant, whereas processing in relation to safety questions can be undertaken in some jurisdictions with a rather different approach. A second, more conceptual yet very interesting

observation can be made, however: individual autonomy can be overridden for solidaristic needs where there is a political will.

4.3.8. Other jurisdictions

While the specifics of data law will of course vary between jurisdictions, many non-EU countries adopt an approach somewhat similar to that of the GDPR. It is not envisaged that these jurisdictional variations will necessarily impede or obstruct the increased use of RWD, but national legislation and regional frameworks must of course be taken into account.

African countries are being called on to sign up to and ratify the African Medical Agency (AMA) Treaty, which is designed to harmonise and accelerate approval of new medicines and vaccines across the continent. The AMA treaty was established in 2019, and by early 2021, 19 countries had signed it, and over half of the 15 countries required to ratify it, had done so. Ultimately, the aim of the treaty is 'to help African countries fight disease outbreaks by ensuring that only high-quality medicines, vaccines, and other health-related supplies reach the market.'[8] By enabling regulatory harmonisation, the AMA and its associated treaties will also facilitate the use of RWD and RWE.

In Brazil, the General Data Protection Law (Lei Geral de Proteção de Dados Pessoais, or LGPD) features key principles of data processing and privacy by design similar to the GDPR. Indeed, the former is described as follows: 'the principle of the purpose of data processing established in the LGPD requires that the purposes of the processing are legitimate, specific, explicit and informed to the data subject. Further processing will only be possible if it is compatible with these purposes.'[9] While a waiver of consent is possible under certain circumstances, 'even if the consent of the data subject for the processing of data by the public authority is eventually waived, in the legally defined cases, such waiver does not exempt the public administration from complying with the other obligations of the LGPD, in particular the general principles and the guarantee of the rights of holders'.[10]

In Canada, the federal Personal Information Protection and Electronic Documents Act (PIPEDA) has governed data use for over two decades, and in addition, some provinces have their own health privacy law or other relevant provincial privacy legislation.

PIPEDA diverges from the GDPR in several ways; PIPEDA does not define personally sensitive information, but medical records are almost always considered sensitive. Notably, PIPEDA applies only to organisations involved in commercial activities and does not apply to public bodies. Furthermore, while the GDPR sets out a number of purposes for legitimate bases for processing data, PIPEDA has a general requirement that organisations 'may only collect, use or disclose personal information for purposes that a reasonable person would consider appropriate'. PIPEDA only requires the data transferring body to ensure protection, while the GDPR also imposes this requirement on the recipient. Whereas PIPEDA places the onus of ensuring comparable protection on organisations carrying out data transfers, the GDPR places that onus on both the exporter and recipient organisations. The GDPR is also stricter on data impact assessments, making them mandatory in certain circumstances, while PIPEDA only recommends them. Finally, 'the GDPR and PIPEDA are inconsistent with respect to the right to erasure, the right to be informed, and the right to data portability'.

In Japan, the Act on the Protection of Personal Information was amended in 2020, and the ethical guidelines for Life Sciences and Medical Research Involving Human Subjects and associated guidance accordingly underwent minor revisions and were published in 2022. According to the Act, 'personal information' means data 'containing a name, date of birth, or other descriptions' or data 'containing an individual identification code … able to identify a specific individual'.

A special category of "'Special care-required personal information' concerns data regarding a person's 'race, creed, social status, medical history, criminal record, fact of having suffered damage by a crime, or other descriptions etc. … of which the handling requires special care so as not to cause unfair discrimination, prejudice or other disadvantages.'".

Similar to the requirements of the GDPR, the Act requires subjects to be told about the use of their data, unless 'it is impossible or requires a disproportionate effort so to do'. Academic institutions are subject to an exception that enables them to use observational personal and clinical data without seeking consent provided that opt-out is possible. In practice, posters in medical centres and information on websites are normally considered sufficient in line with the minimal requirement of 'Guaranteeing opt-out opportunities through disclosure of information'.

Secondary processing of pseudonymised data is only permitted following an institutional ethics committee's approval. Such approval is also required for sharing between institutions.

4.4. Summary

It is clear that RWE is increasingly being used in practice, and we hope to have shown in this chapter that it would indeed be unethical not to increase its use. RCT evidence is still important, but it cannot be generated in some contexts and its focus on selected patients who are often highly unrepresentative of the populations in whom new medicines and other interventions will be used, combined with considerable neglect of some underserved populations such as pregnant women, older patients and ethnic minorities, and the specific issue of the efficacy-effectiveness gap, mean that using RWE to augment RCT evidence is an ethical imperative. Therefore, RWD and RWE have the potential to address health care disparities and promote equity in diverse populations, leading to more inclusive and equitable regulatory decisions.

Given that many treatments are currently prescribed based on old and unrepresentative RCT evidence, this means both that patients may be being prescribed medicines that will not help them, or at least will not help them as much as they and the HCP think, but also that these medicines may cause more harm than predicted. This means that the principles of beneficence (aiming to benefit patients) and nonmaleficence (aiming to not harm them) are threatened by not using RWD; in turn, it means that if HCPs and patients do not know this, then decisions made may be uninformed, threatening autonomy. At a larger scale, the use of unrepresentative data across health systems threatens the principle of justice by distributing resources according to similarly flawed decisions. Equally, of course, from any RWD used, the RWE must be reliable and robust, or decisions made using it will be equally flawed, albeit in a different way from many decisions made using RCT data alone.

Ethical frameworks, guidance, regulations and legislation must be future-proofed to enable RWE to be used in a way that does not violate the autonomy of patients, while also protecting them from the harms that could result from underusing RWD. This chapter is a first step towards this important aim, but the shape and structure of such frameworks will have to be discussed at the societal level. In the COVID-19 pandemic, most members of the public became accustomed to having (some of) their health data used for the greater good. This type of solidarity and greater emphasis on preventing harm and preserving autonomy by ensuring informed decision making about medicines, rather than traditional protection of autonomy by keeping personal data siloed and sealed off, are likely to be key to increasing use of RWE in an ethically robust manner.

4.4.1. An imperative to harmonise

There is an urgent need for principles from the regulators, and for regulators to come together to harmonise the approach taken on data privacy. The lack of guidance at least gives an opportunity for strong guidance to be created now to fill the gaps. What should be the political or philosophical line that is drawn through the guidance?

What can be seen throughout the data protection law is that the legislation has routes that can accommodate different conceptual and political desires. There is a strong rhetorical line that accompanied the implementation of the GDPR towards a conservative reading of the different elements of the law under the desire to ensure individual autonomy. Equally, elements such as the use of the public interest as the legal basis for processing for research purposes; a broad use of informed consent or of compatible processing; and an imaginative use of public notification of data subjects where research in the public interest is being conducted – all these options allow for more research-enabling reading of the legislation for secondary processing of already-gathered personal data in circumstances where, for example, research is being conducted under the approval of RECs, if not under their observation and monitoring. For legislation to fully protect citizens' interests, any potential abuse of personal data must be properly evaluated; robust and effective safeguards should be introduced that avoid such abuse. It is inexcusable to jeopardise citizens' dignity by using ineffective and outdated measures to process personal data. Outmoded measures can create barriers to legitimate and effective data processing that aims to improve and protect citizens' health.

Perhaps COVID-19 is a beginning to a change in the approach. It is increasingly said that the pandemic brought an alignment of incentives for processing personal data. There was a much greater shared interest to use whatever data was available to understand the nature of the virus and to develop vaccines against it. RWD came to the fore, and the pre-pandemic primacy of individual autonomy was relaxed. This is not to say that there were no regulations or safeguards in place. Far from it, the work was conducted under the scrutiny of IRBs and RECs and within the professional integrity of researchers.

The need for robust and joined-up data protection law could not be clearer. RWD offers a huge potential to benefit people. Equally, individuals need protection from breaches of their privacy. Commercial interests cannot be tone deaf to the context within which they seek access to individuals' data; citizens equally cannot be tone deaf to the competing claims they make on society. If commercial interests request altruism from data subjects, they must respond by providing altruistic access to their products and research; if citizens want the benefit of new therapies and pharmaceuticals, they must acknowledge that this requires access to their data.

Chapter 4 – References

1 Seruga B, Templeton AJ, Badillo FE, Ocana A, Amir E, Tannock IF. Under-reporting of harm in clinical trials. *The Lancet Oncology*. 2016;17(5):209-219. https://doi.org/10.1016/S1470-2045(16)00152-2.

2 International Federation of Pharmaceutical Manufacturers & Associations (IFPMA). Patients, product development partnerships and private health sector representatives call for the rapid ratification of the African Medicines Agency Treaty to speed up timelines to vaccines and medicines approval. 2021. (Website accessed 15 April 2023).

3 Templeton AJ, Booth CM, Tannock IF. Informing patients about expected outcomes: the efficacy-effectiveness gap. *Journal of Clinical Oncology*. 2020;38(15):1651-1654. https://doi:10.1200/JCO.19.02035.

4 The Organization for Economic Co-operation and Development (OECD). The OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. 1980. Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data OECD/LEGAL/0188. 2013. (PDF accessed 15 April 2023).

5 Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. *N Engl J Med*. 2019 Nov 14;381(20):1909-1917. https://www.nejm.org/doi/10.1056/NEJMoa1901183.

6 The Organization for Economic Co-operation and Development (OECD). The OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. 1980. Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data OECD/LEGAL/0188. 2013. (PDF accessed 15 April 2023).

7 Compare, for example, Council of Europe (1981) The Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data CETS No. 108 (PDF accessed April 2023) and Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 (Website accessed 15 April 2023) on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

8 Comitê Central de Governança de Dados (Central Data Governance Committee of the Brazilian Government). Guia De Boas Práticaslei Geral De Proteção De Dados (LGPD). August 2020. (PDF accessed 15 April 2023).

9 Faden RR, Beauchamp TL, Kass NE. Informed consent, comparative effectiveness, and learning health care. *New England Journal of Medicine*. 2014;370(8):766-768. http://doi/full/10.1056/NEJMhle1312508.

10 Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials*. 2015;16(1):1-4. https://doi.org/10.1186/s13063-015-1023-4.

CHAPTER 5.
CONCLUSIONS AND FUTURE DIRECTIONS

Over recent years, the RWE field has evolved tremendously and continues to evolve today, driven by recognition and acceptance by regulators, payers and HTA bodies to answer specific research questions, notably during the COVID-19 pandemic. RWE has already been included in various regulatory authorities' approval procedures, reflecting its acceptance and growing importance in evaluating new medicinal products and diagnostics.

Though differences in engagement and process of submission by countries exist, RWE in diverse phases of the product lifecycle has been accepted in many countries. Principles and guidance have been developed by regulators and other stakeholders across the world to support submission of RWE for decision making.

Common stakeholder requirements and expectations are for high quality data and information as well as reliability, accessing and understanding of the information. Continuing the effort of protocol harmonisation and transparency, data quality and integrity framework (including metadata) and interoperability of data, will support standard review of proposed evidence plan including RWD, as well as the generated RWE. These activities will strengthen the grounds for RWE acceptance and support the development of evolving technologies and methods, including artificial intelligence and also potentially open access to different sources of data, e.g. health care sensor for remote monitoring.

RWE could and should be considered, if appropriate, because the strategy for addressing evidence gaps should cover all types of evidence generation, whether from a clinical trial or an OS, and should only be based on the research question of interest. If RWE is fit for purpose, it is best to engage early with regulators to facilitate discussion on the evidence plan as understanding of the RWD source will be critical in this discussion.

RWD has been used to evaluate the safety of medicinal products for regulatory decision making for decades, and more recently for effectiveness as well. While the common misunderstanding is that RWD includes only EHR, the scope is much broader. It also includes sources such as SRSs and surveys. Importantly, RWD sources can be categorised differently, and the categorisation in this report may be different from that in other documents. The report uses categorisation based on the principle of how different RWD sources can produce different types of RWE for different purposes to address safety and effectiveness of medicinal products.

Each RWD source has its strengths and limitations, and the source may be useful for certain safety and effectiveness purposes, but not for others. Survey data sources are very useful for estimating the burden of diseases, but they are not the most appropriate for exploring association between medicinal products and outcomes, which require follow-up information. Scientific evaluation of the fitness of a RWD source for a given study is essential in choosing a data source.

The rapid development in new technologies has very quickly resulted in new RWD sources with large volumes. Although the current use of these emerging sources is limited because of their complexity, which require a new set of methods, they have a great potential to be key RWD sources in the context of regulatory decision making in the future.

The key scientific considerations regarding the design and analysis of studies that generate RWE have been discussed. The specification of a clear question reflects both the regulatory and clinical context. The assessment of health care data resources as fit for purpose is specific to this

question and includes: detailed assessment of the extent of missing data; reliability and validity of key constructs; and integrity of the data including transformations. Study design decisions (e.g. selection of the comparator; identification of the population of interest; and timing of exposure, outcome and confounder measures) affects the validity and generalisability of the study results, and thus are essential to the generation of fit-for-purpose RWE.

Emulating a RCT for designing studies using RWD is an approach that seeks to address the limitations of OSs in evaluating the safety and effectiveness of medical interventions. Advantages have been described, but most importantly, they clarify thinking while making crucial design decisions such as inclusion criteria, duration of follow-up, and study endpoints, and reduce the potential for introducing error. Shortcomings in the study design are often difficult, at best, to overcome in the analysis.

There has been a misunderstanding that only observational (non-interventional) studies use RWD to produce RWE — a dichotomy between controlled trials and observational studies.[1] In fact, controlled trials can also use RWD studies and produce RWE.[2] For example, RCTs where the primary outcomes are assessed using RWD generate RWE (pragmatic trials). Externally controlled trials also generate RWE when the control arms are constructed from RWD sources.[3]

The statistical analysis plan should be aligned with the research question and address potential sources of bias due to confounding, measurement error, and selection of participants. Consideration should be given to handling variables including competing risk events and time-dependent variables, and to approach missing data. While in clinical trials comparability among the treatment arms is achieved via randomisation, in RWD studies it can be achieved, among other approaches, by addressing confounders. Statistical methods to improve comparability (e.g. matching and adjusted analysis) have been discussed. In addition to the primary analysis, it is necessary to conduct additional sensitivity analyses to quantify the robustness of the main results to violations of assumptions, plausible degrees of measurement error in key variables, and alternative choices for parameters in the study design (e.g. grace periods and handling of treatment changes during follow-up).

Protocol registration, transparent reporting, and responsible communication of results are all important components of establishing reliable RWE for regulatory decision making.

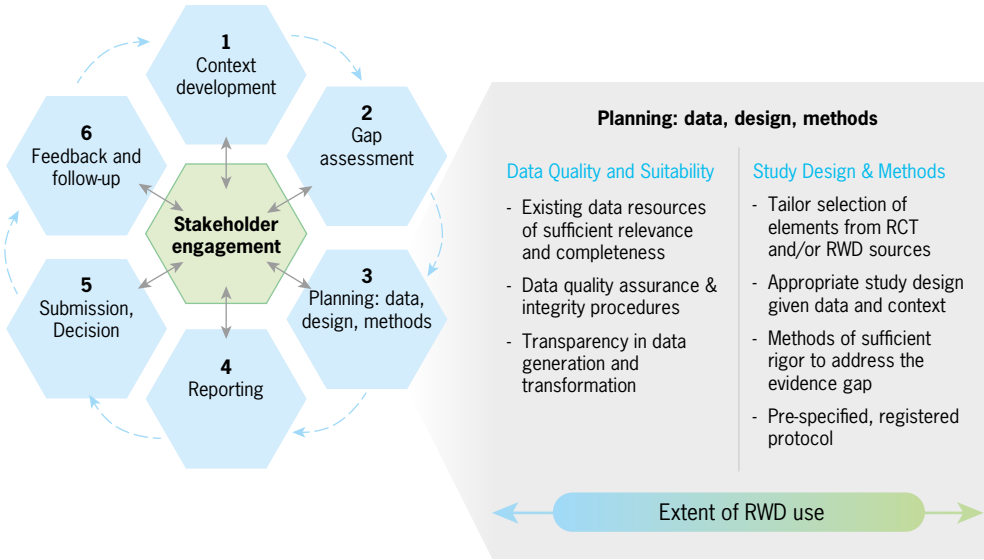
The process of RWE generation described in this report is shown in **Figure 5**. The first step is the articulation of the context. What type of evidence is needed: safety, effectiveness or both? Is the evidence needed to fulfil regulatory requirements or to inform decision making in clinical practice? Who are the stakeholders: regulators, HTAs, or payers?

After the context articulation, the next step is gap assessment to evaluate if such evidence exists and if there is a need to generate new evidence. The results of the gap assessment inform the next step: the identification of the most appropriate RWD, designs, and methods for the study to generate the needed RWE.

The results of the study must be interpreted in the right context for effective submissions and for the right decision making and application. Stakeholders play critical roles along the process to make sure that the studies are planned and performed appropriately with the right context and purpose. Feedback from stakeholders may require follow-up actions, including additional analysis and new studies, and the cycle continues.

Figure 5. Evidence generation cycle

Source: CIOMS Working Group XIII



In terms of the ethical framework, several fundamental questions about data-sharing norms must also be considered, particularly the nature of privacy rights, and how far informed consent is required for the re-use of personal data in different settings from where it was initially gathered.

The current standards and expectations are built on a series of normative assumptions, and these assumptions have been opened up for discussion to create space in the normative discourse to explore the scientific proposals for change.

In the context of efficacy-effectiveness gap, if patients are being given inaccurate information about the potential benefits and risks of treatments for them, then some decisions that use that information may be being made without valid informed consent. Increasing use of RWE is one important way to fill the efficacy-effectiveness gap and complement the evidence from RCTs.

RWD is increasingly used in practice, and this often takes place without any ethical or legal framework specific to the use of RWD.

There is a strong argument that the processing of RWD only works where data subjects have trust and confidence in the institutions and individuals who process data that relate to them; therefore, a strong personal data protection regime is essential to the acceptance and operation of RWD processing.

Further efforts are needed on issues regarding compatible processing of RWD in the absence of consent or where data were gathered to form a patient record.

This report has discussed the role of RWD/RWE in health-related regulatory decision making along the medicinal product's lifecycle and the needs of the different stakeholders, the available data sources, the key scientific considerations, as well as the ethics and governance perspectives. More work remains to be done to globally harmonise practices and guidance for using RWD and RWE for regulatory decision making. We hope this report furthers the discussion to this end and contributes to the future development of RWD/RWE in regulatory decision making, thereby helping to maximise the benefits they can bring to public health.

Chapter 5 – References

1 Concato J, Stein P, Dal Pan GJ, Ball R, et al. Randomized, observational, interventional, and real-world-What's in a name? *Pharmacoepidemiol Drug Saf*. 2020 Nov;29(11):1514-1517. https://doi.org/10.1002/pds.5123.

2 Rahman M, Dal Pan G, Stein P, Levenson M, et al. When can real-world data generate real-world evidence? *Pharmacoepidemiol Drug Saf*. 2024 Jan;33(1):e5715. https://doi.org/10.1002/pds.5715.

3 Rahman M, Dal Pan G, Stein P, Levenson M, et al. When can real-world data generate real-world evidence? *Pharmacoepidemiol Drug Saf*. 2024 Jan;33(1):e5715. https://doi.org/10.1002/pds.5715.

APPENDIX 1.
CASE STUDIES

These case studies complement the chapters in this report; they are not intended for use as guidance. We encourage readers to follow local guiding principles and regulatory guidelines pertaining to RWD and RWE where available.

A. Fosdenopterin approved for treatment of a rare, genetic disease with external control data from a natural history disease study

Topic	Summary Information
Rationale	This case study demonstrates the use of data on natural history of disease as external controls in comparison to single-arm trial data, constituting an adequate and well-controlled study in support of assessment of treatment effectiveness.
Study question	Do patients treated with fosdenopterin show an improved survival outcome compared to untreated patients in a natural history disease study?
Medicinal product	One fosdenopterin product, a synthetic cPMP, was approved in 2021 by the US FDA for the treatment of molybdenum cofactor deficiency (MoCD) Type A. No pharmaceutical treatment was approved before fosdenopterin for this rare and fatal disease. Patients with MoCD Type A do not have enough cPMP, a substance needed to synthesise molybdenum cofactor.
Indication/Disease treated	MoCD Type A is a rare, neurodegenerative, autosomal recessive disease with an estimated US prevalence of approximately 50 patients, all under 10 years of age. It affects the central nervous system, leading to intractable seizures, metabolic acidosis, failure to thrive, feeding difficulties, axial hypotonia and death in the first years of life mainly from infection.
Stage of the medicinal product development lifecycle	The natural history of disease study was conducted during pre-marketing clinical research.
Where were the study protocols registered?	Unknown

Topic	Summary Information
RWD study design and results	The adequate and well controlled investigation consisted of a comparison of overall survival in 13 patients with MoCD Type A who were treated with fosdenopterin or rcPMP (a recombinantly produced version of the drug with the same active moiety and same biologic activity) to that of an untreated natural history cohort of 18 patients with MoCD Type A who were genotype-matched to the treated patients. The natural history of disease study was a combined retrospective and prospective, non-interventional study collecting data on untreated patients with MoCD Type A. Treated patients showed a significant improvement of overall survival compared to the untreated control patients.
How did the involvement of RWD/RWE in the study affect the study design at the outset?	Clinical trials were designed as single-arm trials at the outset due to the rarity of the disease and the known, strong genotype-phenotype correlation. The natural history of disease study was conducted to provide comparisons to the treated patients in the trials.
What were the data sources used and why were they chosen?	RWE came from a combined retrospective and prospective, non-interventional, natural history of disease study collecting data on untreated patients with MoCD Type A in academic centres in 14 countries.
What were the data analysis methods used? Why were they chosen?	Data analysis used the log-rank test to compare treated and natural history control patients, and Kaplan–Meier plots and methods to estimate survival parameters for each group. Additionally, the statistical analysis plan specified analysing overall survival using the Cox proportional hazards model by regressing survival on an indicator variable denoting treatment status.
What legal data protection requirements had to be met in the countries you were working in?	They seem to be country specific.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	It is part of NDA submitted to the US FDA in support of the effectiveness and safety evaluation. The comparison of overall survival in patients treated with fosdenopterin to that in an untreated, natural history cohort of patients who were genotype-matched to the treated patients constitutes an adequate and well controlled investigation in the context of the very rare disease that was rapidly fatal with no other therapies known to improve survival. The efficacy data were adequate to support a conclusion that fosdenopterin provides a survival benefit in patients with MoCD Type A.

Topic	Summary Information
Conclusion	When designed and conducted properly, external controls from RWD sources can provide RWE that support of regulatory decision making. The strengths of the natural history data lie in the use of a reliable and objective endpoint (mortality) and that the external control patients were genotype matched to the treated patients. The confirmatory evidence includes biomarker data results which provide assurance. The benefits of fosdenopterin outweigh its risks when used according to the product labelling.
Published reference	Kang, C. Fosdenopterin: first approval. *Drugs* 81, 953–956 (2021). https://doi.org/10.1007/s40265-021-01520-2.

B. Comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with different COVID-19 vaccines in an international network cohort study

Topic	Summary Information
Rationale	The study aimed to quantify the comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with adenovirus-based COVID-19 vaccines versus mRNA-based COVID-19 vaccines to understand the potential risks of some available vaccines compared with each other. This case study is based on the application of the OMOP CDM techniques. The OMOP CDM is a CDM that provides a standardised way to represent and organise observational health care data from disparate sources, enabling data harmonisation and facilitating data sharing and collaboration across different healthcare systems and research institutions (see Chapter 1).
Study question	Are risks of thrombosis with thrombocytopenia syndrome or thromboembolic events in adenovirus based versus mRNA based COVID-19 vaccines different?
Medicinal product	Four COVID-19 vaccines were included: ▶ ChAdOx1-S (Vaxzevria, AstraZeneca); ▶ BNT162b2 (Comirnaty, BioNTech and Pfizer); ▶ mRNA-1273 (Spikevax, Moderna); and ▶ Ad26.COV2.S (Jcovden, Janssen). The ChAdOx1-S and the Ad26.COV2.S vaccines use weakened versions of adenoviruses. The adenovirus is modified to carry the genetic code for the spike protein found on the SARS-CoV-2 virus, which causes COVID-19. When the vaccine is given, the adenovirus delivers the spike protein genetic code to cells in the body, causing them to produce the spike protein. The immune system then recognises the spike protein as foreign and produces antibodies to attack it. The BNT162b2 and the mRNA-1273 vaccines are messenger RNA (mRNA) vaccines. This type of vaccine uses a small piece of genetic material (the mRNA) that codes for SARS-CoV-2 spike protein. When the vaccine is given, the mRNA enters cells in the body and instructs them to produce the spike protein.

Topic	Summary Information
Indication/Disease treated	COVID-19 is a highly infectious disease caused by the coronavirus SARS-CoV-2. The virus was first identified in Wuhan, in the People's Republic of China, in December 2019, and has since spread rapidly to become a global pandemic. Vaccines against COVID-19 have been developed and are being distributed around the world, with the aim of preventing severe illness, hospitalisation, and death from the disease.
Where were the study protocols registered?	ENCePP database with the id. EUPAS44469.
Stage of the medicinal product development lifecycle	The study took place at post-marketing stage of the vaccines.
How did the involvement of RWD/RWE in the study affect the study design at the outset?	This study was the first multinational analysis of the comparative safety of adenovirus-based COVID-19 vaccines compared with mRNA-based ones, using data routinely obtained in diverse databases in several countries, and at the same time based on the use of common standards and data model. The OMOP CDM allowed the study to be run by each site with common analytical code.
What were the data sources used and why were they chosen?	The study used datasets from five European countries (France, Germany, the Netherlands, Spain and the UK) and two datasets from the US including more than 3 million patients. All these databases are in OMOP CDM format. The datasets included electronic health care records collected from patients registered with general practices, primary care records databases, hospital discharge data and medical claims. The datasets were anonymised to protect patient privacy.

Topic	Summary Information
What were the data analysis methods used? Why were they chosen?	The study used descriptive statistics to report the baseline characteristics for each cohort. Propensity scores were calculated for each pair of vaccines being compared, and patients were matched using greedy matching.[i] The study used three diagnostic tools to evaluate measured confounding, statistical power, and unmeasured confounding. Poisson regression was used to calculate the incidence rate ratio and 95% confidence intervals of outcomes according to the target and comparator vaccinations. Empirical calibration was used to account for residual systematic error due to potential unobserved confounding. Finally, random effect meta-analysis was conducted to pool results across databases. OMOP CDM has been widely adopted and validated for active safety surveillance research and comparative effectiveness studies, facilitating large-scale, multi-institutional research projects. The OMOP CDM enables researchers to perform more comprehensive analyses of RWE, which can inform clinical practice and policy decision making. One limitation is the need for data mapping and terminology standardisation, which can be resource-intensive and time-consuming. Another limitation is the potential for bias and confounding in observational data, which can affect the validity and reliability of research findings. Additionally, the quality and completeness of data can vary across different sources, which can impact the generalisability and usefulness of research findings.
What legal data protection requirements had to be met in the countries you were working in?	The study protocol for this research was approved by the independent scientific advisory committee for UK's Medicine and Healthcare Products Regulatory Agency database research (protocol No 21_000641). Informed consent of individual patients was not required as anonymised information was obtained from medical records.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	The study was funded by EMA. EMA 2017/09/PE – Association between thrombosis with thrombocytopenia syndrome (TTS) or thromboembolic events, and COVID-19 vaccines. Procurement procedure no. EMA/2017/09/PE (Lot 3) The use of RWE can help to improve the efficiency and speed of regulatory decision making and can provide important insights into the real-world benefits and risks of a treatment. However, it is important to ensure that the RWE is of high quality and that appropriate methods are used to account for potential biases and confounding factors.

[i] ***Greedy matching***
Greedy caliper matching is a popular method used in PS matching. This method orders the treated subjects, and the first treated subject is randomly matched to an untreated (or alternatively treated) subject with a PS that is within a predefined caliper width. The initial ordering of subjects is often done randomly but may also be based on a subject's PS or other parameters. In addition to caliper matching, nearest neighbor (NN) caliper matching is often used, where the treated subject is matched to an untreated subject that has the closest propensity score within the caliper. Both methods do not consider that the untreated subject can potentially form a better pair with another treated subject that is further down the line; hence they are "greedy" algorithms. Because of this, both methods are dependent on the random order in which the treated subjects are placed, if patients are not ordered based on their PS. In addition, caliper matching is also dependent on which untreated patient within the caliper is randomly matched.
***Source:* Austin PC. Propensity-score matching in the cardiovascular surgery literature from 2004 to 2006: A systematic review and suggestions for improvement. J Thorac Cardiovasc Surg; Mosby. 2007;134:1128-1135.e3.** [PubMed] [Google Scholar]

Topic	Summary Information
Conclusion	This study provides a key context on the complications in unvaccinated participants suffering from COVID-19, showing a remarkable increase in the risk of some outcomes in these patients, such as pulmonary embolism, disseminated intravascular coagulation, or myocarditis. This study has important strengths, including the use of a cohort study with active comparators and replication of the exact same analysis across different databases using the OMOP CDM. This study has some limitations due to heterogeneity across data sources. Information bias due to outcome ascertainment was likely present, and the study was susceptible to unmeasured confounders.
Published reference	X Li, E Burn, T Duarte-Salles, C Yin, et al. Comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with different COVID-19 vaccines: international network cohort study from five European countries and the US. *BMJ* 2022; 379: e071594. https://doi.org/10.1136/bmj-2022-071594.

C. Contextualising adverse events of special interest to characterise the baseline incidence rates in 24 million patients with COVID-19 across 26 databases: a multinational retrospective cohort study

Topic	Summary Information
Rationale	The objective of this study was to estimate the incidence rates of several adverse events of special interest (AESI) related to vaccination among individuals with COVID-19, compared to the rates in the general population before the pandemic. It should be noted that some AESIs may not only be potentially linked to COVID-19 vaccines but also to SARS-CoV-2 infection itself. Therefore, to evaluate the benefits and risks of COVID-19 vaccines properly, it is crucial to consider the expected occurrence rates of these events in individuals with COVID-19. To address this issue, the OHDSI community conducted a network study using data from 26 databases across 11 countries. This case study is based on the use of OMOP CDM standards and techniques. The OMOP CDM is a standard data model for organising and analysing observational health data, including EHRs, insurance claims, and other healthcare administrative data. It was developed by the OHDSI community to enable the sharing and analysis of large-scale health data across different databases and research studies (see Chapter 2).
Study question	What is the evidence on the occurrence of AESI after COVID-19 infection rather than after vaccination?
Medicinal product	This study focused on the study of AESI after COVID-19 disease. AESI are specific and significant adverse events potentially related to a particular medical intervention, such as a medicine or vaccine. The identification and monitoring of AESIs is an important part of drug safety surveillance and regulatory decision making. By closely monitoring AESIs, regulatory authorities can ensure that medical interventions are safe and effective for patients.

Topic	Summary Information
Indication/Disease treated	COVID-19 is a respiratory disease caused by the SARS-CoV-2 virus that was first reported in Wuhan, in People's Republic of China, in December 2019 and has since spread globally. The disease is primarily transmitted through respiratory droplets released when an infected person talks, coughs or sneezes. Common symptoms of COVID-19 include fever, cough, and fatigue, while more severe symptoms such as shortness of breath, pneumonia, and mortality also occur. Disease severity varies by age and underlying health conditions. As of February 2024, the number of confirmed COVID-19 cases worldwide exceeded 774 million, resulting in over 7 million deaths.[1] In contrast, AESI associated with COVID-19 vaccines are generally rare, with the most common AESIs being mild and temporary, such as pain at the injection site or fever. However, serious AESIs have been reported in some cases, including blood clotting disorders and myocarditis, and the risk of severe AESI following COVID-19 vaccination varies depending on the age, sex, and underlying health conditions. It is important to note that assessing the relationship between COVID-19 vaccines and AESI can be complicated because some AESIs may be associated with COVID-19. The AESI included in the study are: Guillain-Barré syndrome, facial nerve (Bell's) palsy, anaphylaxis, encephalomyelitis, narcolepsy, appendicitis, non-haemorrhagic stroke, haemorrhagic stroke, acute myocardial infarction, myocarditis and pericarditis, deep vein thrombosis, pulmonary embolism, disseminated intravascular coagulation, immune thrombocytopenia, transverse myelitis, and the co-occurrence of thrombosis with thrombocytopenia.
Where were the study protocols registered?	The complete information is publicly available at the OHDSI initiative website: https://ohdsi-studies.github.io/Covid19SubjectsAesiIncidenceRate/Protocol.html.
Stage of the medicinal product development lifecycle	The study took place at post-marketing stage and was focused on the detection of conditions included under the definition of AESI but related to COVID-19 disease and not in specific medicinal products.

Topic	Summary Information
How did the involvement of RWD/RWE in the study affect the study design at the outset?	The study used data routinely obtained from diverse databases in several countries, which had in common the use of the same standards and data model. The OMOP CDM allowed the study to be run by each site using the same analytical codes and bio-informatic tools. The total number of participants included in all databases was 945,520,607. OSs have investigated the incidence rates of AESI among patients with COVID-19 and those vaccinated against COVID-19. To accurately assess the benefit–risk of COVID-19 vaccines, it is essential to carefully analyse the available epidemiological data on both COVID-19 disease and vaccination. Such analysis should take into consideration potential confounding or intermediating factors that may affect the observed association between vaccines and AESI.
What were the data sources used and why were they chosen?	The study included 23,840,986 patients with COVID-19 from 26 databases representing a diverse set of care settings from North America, Europe, and Asia including the following 11 countries: Belgium, Estonia, France, Germany, Japan, the Netherlands, Serbia, Spain, Turkey, the UK, and the US. All these databases were harmonised and standardised in the OMOP CDM format. The datasets included electronic health care records collected from patients registered with general practices, primary care records databases, hospital discharge data, and medical claims. The datasets were anonymised to protect patient privacy.
What were the data analysis methods used? Why were they chosen?	Incidence rates were calculated by dividing the total number of events by person-time at risk and were stratified by age and sex subgroups for each database. The rates were pooled across the databases using a random effects meta-analysis, and indirect standardisation was used to account for differences between age subgroups and sex distribution in the COVID-19 cohort and the pre-pandemic background population. The study also used negative control outcomes to evaluate potential bias in incidence ratio estimates. The meta-analytic rates were classified according to the CIOMS thresholds: very common (≥ 10%), common (> 1% to < 10%), uncommon (≥ 0·1% to < 1%), rare (≥ 0·01% to < 0·1%), and very rare (< 0·01%). Mapping all the databases to the OMOP CDM standards was used. OMOP CDM has been widely adopted and validated for active safety surveillance research and comparative effectiveness studies, facilitating large-scale, multi-institutional research projects. One limitation is the need for data mapping and terminology standardisation, which can be resource-intensive and time-consuming. In addition, EHR databases may not capture all medical events that occur outside the participating health system, leading to incomplete information. To reduce the impact of incomplete data, the study only included patients who had at least one year of continuous observation. However, defining continuous observation can be problematic when working with diverse databases.

Topic	Summary Information
What legal data protection requirements had to be met in the countries you were working in?	Informed consent from individual patients was not required as information from the different clinical databases was anonymised. The study protocol was approved by the IRB committees of the participant databases. In addition, the New England Institutional Review Board has determined that some databases are exempt from study-specific IRB.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	The study was partially funded by the EHDEN from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 806968. The JU receives support from the EU's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations (EFPIA). RWE has the potential to enhance the efficiency and speed of regulatory decision-making processes and offer valuable insights into the actual benefits and risks associated with treatment. Nonetheless, it is critical to guarantee the quality of RWE and implement appropriate techniques to adjust for potential biases and confounding variables.
Conclusion	The study suggests that COVID-19 disease itself must be considered when assessing the relationship between COVID-19 vaccines and AESI, as it may confound or mediate the observed association. When conducting OSs on this subject, controlling for COVID-19 is crucial. The strength of this study lies in its use of a large number of patients and databases from different regions, enabling a comprehensive assessment of AESI incidence rates among patients with COVID-19. At the moment of publication, it is the largest study about COVID-19, including about 24 million people with COVID-19 and over 945 million general population participants, from 26 data sources across three continents. Regarding its limitations, the study did not differentiate between COVID-19 variants or consider recurrent COVID-19, limiting its ability to compare the AESI incidence rates between different variants or patients with multiple infections.
Published reference	Voss EA, Shoaibi A, Yin Hui Lai L, Blacketer C, et al. Contextualising adverse events of special interest to characterise the baseline incidence rates in 24 million patients with COVID-19 across 26 databases: a multinational retrospective cohort study. *eClinicalMedicine* 2023;58:101932. https://doi.org/10.1016/j.eclinm.2023.101932.

D. N-Nitrosodimethylamine (NDMA)-contaminated valsartan and the risk of cancer

Topic	Summary Information
Rationale	The study provides an example of RWD use in the post-authorisation setting of drug safety analyses. It is an example of how RWD from statutory health insurance can be used to examine urgent drug safety questions with pharmacoepidemiological methods. The immediate recall of all potentially *N*-nitrosodimethylamine (NDMA)-contaminated valsartan medicines by regulatory authorities worldwide was necessary to protect public health. The detection of different nitrosamine impurities in medicines since 2018 led to the introduction of a new threshold by the EMA.
Study question	Is there an association between filled prescriptions of potentially NDMA-contaminated valsartan medicines and cancer risk in comparison with non-contaminated valsartan in routine care in Germany?
Medicinal product	Valsartan is an angiotensin II receptor antagonist, typically administered as tablets. There are different MAHs. In 2018, NDMA was detected in the valsartan active substance manufactured by one particular sponsor. Preparations containing the contaminated valsartan were withdrawn from the market by regulatory agencies across the world.
Indication/Disease treated	The angiotensin II receptor antagonist valsartan is used predominantly to treat hypertension and heart failure. Valsartan blocks the actions of angiotensin II, which include constricting blood vessels and activating aldosterone, to reduce blood pressure. The drug binds to angiotensin type I receptors (AT1), working as an antagonist. This mechanism of action is different than that of the ACE inhibitors, which block the conversion of angiotensin I to angiotensin II.
Where were the study protocols registered?	N/A as this was a research paper.
Stage of the medicinal product development lifecycle	Post-marketing
How did the involvement of RWD/RWE in the study affect the study design at the outset?	In 2018, NDMA was detected in the valsartan active substance but the contamination of valsartan seemed to be the result of a change in the manufacturing process in 2012. Therefore, a retrospective cohort study was initiated.

Topic	Summary Information
What were the data sources used and why were they chosen?	The study is based on longitudinal routine data from a large German statutory health insurance provider. On average, nearly 25 million persons were insured by the insurance provider each year during the study period. Furthermore, MAHs provided batch-related data on all valsartan medicines for the study period. This included information on which batches were manufactured using the potentially contaminated valsartan and how many packages of these medicines were sold. The long duration (2009–2017) and the large sample size (780 871 patients were included for analyses) were important criteria for being able to observe the association of NDMA contamination with the risk of cancer.
What were the data analysis methods used? Why were they chosen?	We used Cox regression models with time-varying variables and with adjustment for potential influencing factors to calculate hazard ratios (HR) for cancer overall and for several individual cancer types.
What legal data protection requirements had to be met in the countries you were working in?	The routine data used for the study cannot be shared with or transmitted to third parties due to legal restrictions.
What did you change (if anything) to be in line with ethical considerations?	The study protocol is in line with ethical considerations.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	With our study regulatory authorities worldwide receive information to assess the public health impact of NDMA contamination in valsartan medcines. The study is an example of how to use pharmacoepidemiological methods and RWD to examine urgent questions of drug safety.
Conclusion	The conclusion of the study included that careful monitoring of potential further effects of NDMA-contaminated valsartan after longer periods is advisable.
Published reference	Gomm W, Röthlein C, Schüssel K, Brückner G, et al. N-Nitrosodimethylamine-contaminated valsartan and the risk of cancer—a longitudinal cohort study based on German health insurance data. *Dtsch Arztebl Int* 2021; 118: 357–62. https://doi.org/10.3238/arztebl.m2021.0129.

E. Cardiovascular risk of urate-lowering medicines: a study using the National Database of Health Insurance Claims and Specific Health Check-ups of Japan

Topic	Summary Information
Rationale	This case study considered how the risk of cardiovascular death in patients with gout was higher in a group taking febuxostat than in a group taking allopurinol in the CARES trial (Cardiovascular Safety of Febuxostat and Allopurinol in Participants With Gout and Cardiovascular Comorbidities); however, the extrapolation of these results to Japan remained unclear. The specific aim of this study was to compare the risk of cardiovascular events associated with febuxostat and topiroxostat with that associated with allopurinol in Japan. The primary outcome of this study was the occurrence of cardiovascular events, including acute coronary syndrome, cerebral infarction, and cerebral haemorrhage, during the follow-up period. Cardiovascular death was set as the secondary outcomes in addition to an individual component of the primary outcome. See Section 1.6.3.
Study question	Is the risk of cardiovascular events associated with urate-lowering medicines in Japan? The primary focus of the study was on the risk of febuxostat and topiroxostat when compared with allopurinol in Japan.
Medicinal product	Febuxostat or topiroxostat for exposure groups, allopurinol for the control group, and benzbromarone for the secondary control group. Febuxostat, topiroxostat and allopurinol reduce serum uric acid through an inhibitory action of xanthine oxidase. Benzbromarone promotes uric acid excretion by inhibiting uric acid reabsorption in the tubules.
Indication/Disease treated	Febuxostat Indication: (1) Gout, hyperuricemia, and (2) Hyperuricemia associated with chemotherapy.
Stage of the medicinal product development lifecycle	Post-marketing
Where were the study protocols registered?	The protocols were registered with the PMDA.
How did the involvement of RWD/RWE in the study affect the study design at the outset?	The large size of claims data with a long follow-up period allowed enough sample size to detect relatively rare cardiovascular events and to quantitatively compare risk among different medicines.

Topic	Summary Information
What were the data sources used and why were they chosen?	Data from the NDB were used in this study because (1) the NDB is the largest database managed by the MHLW, collecting information on nation-wide medical claims from hospitals, clinics, pharmacies, and dental clinics; and (2) the NDB ensures a long follow-up period e.g. from hospitals where patients undergo treatment.
What were the data analysis methods used? Why were they chosen?	The incidence rates of outcomes (primary and secondary outcomes) in each group were calculated, followed by calculating the incidence rate ratio of the exposure groups to the control group (allopurinol). Crude and adjusted hazard ratios were also estimated using the Cox proportional hazards model with the adjusted factors for assuring appropriate comparability of groups.
What legal data protection requirements had to be met in the countries you were working in?	The data in NDB was anonymised for protecting personal information, and did not include personal information such as patient names, addresses, or names of medical personnel. Since NDB is operated by MHLW in accordance with the law, it is not necessary to obtain consent from patients for the collection of their medical information. For promoting the appropriate use of medical information, the study plan and results for publication, etc., are required to comply with the user guideline of NDB.
What did you change (if anything) to be in line with ethical considerations?	As this study was conducted as an official activity of the PMDA under the PMDA law, it was not subject to review by IRBs.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	The PMDA conducted a safety assessment of the risk of febuxostat and topiroxostat based on this study's results and other available data, including spontaneous adverse drug reaction reports, literature, and the results of the FAST trial (Febuxostat versus Allopurinol Streamlined Trial), and concluded that no additional regulatory actions are currently warranted.
Conclusion	No increased cardiovascular risk was observed with febuxostat or topiroxostat when compared with allopurinol in patients with hyperuricemia in Japan. (The adjusted hazard ratios for the cardiovascular risk were 0.97 (95% confidence interval (CI): 0.95–0.98) for febuxostat and 0.84 (95% CI: 0.78–0.90) for topiroxostat groups). This is the first quantitative assessment of the risk of cardiovascular events associated with febuxostat and topiroxostat when compared with allopurinol in Japan.
Published reference	Sawada S, Kajiyama K, Shida H, et al. Cardiovascular risk of urate-lowering drugs: A study using the National Database of Health Insurance Claims and Specific Health Checkups of Japan. *Clin Transl Sci.* 2023; 16: 206–215. https://doi.org/10.1111/cts.13439.

F. Nested case-control study utilising MID-NET® on thrombocytopenia associated with pegfilgrastim in patients treated with antineoplastic agents

Topic	Summary Information
Rationale	This case study was about investigating the association between human granulocyte colony-stimulating factors (G-CSF) preparations (filgrastim, lenograstim, nartograstim, and pegfilgrastim) available in Japan and thrombocytopenia in patients treated with antineoplastic agents. A nested case-control study was conducted using the MID-NET® with the cohort of the Japanese population taking antineoplastic agents. MID-NET® stores electronic medical records, administrative claims data, and diagnosis procedure combination data of about 5.3 million patients (as of December 2020) in cooperation with 10 health care organisations, including 23 university hospitals or regional core hospitals. See Section 1.6.3.
Study question	Do G-CSF preparations cause thrombocytopenia in patients treated with antineoplastic agents?
Medicinal product	G-CSF preparations (filgrastim, lenograstim, nartograstim, and pegfilgrastim) are human granulocyte colony-stimulating factors. These G-CSF medicines promote the bone marrow to produce white blood cells.
Indication/Disease treated	G-CSF medicines are used for managing neutropenia caused by antineoplastic agents.
Where were the study protocols registered?	The protocols were registered with the PMDA.
Stage of the medicinal product development lifecycle	Post-marketing
How did the involvement of RWD/RWE in the study affect the study design at the outset?	MID-NET® included laboratory test results examined in clinical practice. Thus, platelet count data, which were an appropriate indicator for thrombocytopenia, were used for this study. These data allowed more objective detection of target events (study outcome).
What were the data sources used and why were they chosen?	Data from MID-NET®, a reliable and valuable database in Japan, were used for analysis in this study. In this database, platelet count data, which are an appropriate indicator for thrombocytopenia, are available for analysis. In addition, the outcome of this study (thrombocytopenia after administration of G-CSF preparations during the treatment period with antineoplastic agents) can be obtained in the same hospital.

Topic	Summary Information
What were the data analysis methods used? Why were they chosen?	A nested case-control design was selected to account for many covariates just prior to the occurrence of thrombocytopenia, such as type of antineoplastic agent and its treatment duration, comorbidity, and co-prescribed medicines. To evaluate the association between the use of G-CSF preparations and thrombocytopenia, conditional logistic regression analysis considering with matching factors was conducted to estimate crude odds ratios (ORs) and adjusted ORs (aOR) with adjustment for radiological therapy. Similar analysis was conducted on each medicine in the detailed analysis.
What legal data protection requirements had to be met in the countries you were working in?	The data in MID-NET® is anonymised for protecting personal information, and does not include information such as patient names, addresses, or names of medical personnel. For promoting the appropriate use of medical information, the study plan and results for publication, etc., are required to comply with the user guideline of MID-NET®.
What did you change (if anything) to be in line with ethical considerations?	As this study was conducted as an official activity of the PMDA under the PMDA law, it was not subject to review by IRBs.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	The PMDA conducted a safety assessment on the risk of thrombocytopenia in association with G-CSF preparations based on case reports and related literature as well as the results from this study. In March 2020, the PMDA announced a revision of the package insert of pegfilgrastim to inform on G-CSF-induced thrombocytopenia.
Conclusion	A significantly increased risk of thrombocytopenia associated with pegfilgrastim was identified (aOR: 7.4 95% CI: 2.0–28.1). More attention to thrombocytopenia may be necessary during treatment with pegfilgrastim. This finding was the key evidence for the PMDA regulatory safety action of revising the label (package insert) of pegfilgrastim.
Published reference	Kajiyama K, Ishiguro C, Ando T, et al. Nested case-control study utilising MID-NET® on thrombocytopenia associated with pegfilgrastim in patients treated with antineoplastic agents. *Clin Pharmacol Ther*. 2021; 110: 473–479. https://doi.org/10.1002/cpt.2263.

G. Post-authorisation safety studies for CAR (chimeric antigen receptor) T-cell products

Topic	Summary Information
Rationale	The case study is a sample of regulatory use of patient registry data for PASS of chimeric antigen receptor (CAR) T-cell products. It is a RWD study to collect information on future expansion of indications for the same product or different products targeting different antigens. International harmonisation of survey items is being done through international collaboration of academic groups that had been conducting work with highly inclusive patient registries.
Study question	To determine the safety and efficacy of CART-cell therapy products post-authorisation in real-world.
Medicinal product	CAR T-cell therapy is an innovative form of immunotherapy that has gained significant attention in recent years. It involves genetically modifying T-cells to express a specific receptor, known as a CAR, which allows the infused cells to recognise and bind to specific antigens on cancer cells and to attack the cancer cells effectively by using patients' own immune reaction. CAR T-cell therapy using CD19-specific CAR has been reported to have dramatic effects on B-cell tumours, which led to the approval of the first commercial product for relapsed/refractory B-cell acute lymphoblastic leukaemia (ALL) and relapsed/refractory diffuse large B-cell lymphoma (DLBCL) in 2017 in the US, in 2018 in the EU, and in 2019 in Japan. CAR T-cell products targeting B-cell maturation antigen (BCMA) were approved for the treatment of multiple myeloma in 2021 in the US and in the EU, and 2022 in Japan. Current approved products are autologous CAR T-cells, which are processed by using the patient's harvested T cells. *Tisagenlecleucel:* CD19 directed CAR T-cell product. The first CAR T-cell therapy product was approved by the US FDA in August 2017, by EMA in August 2018, and March 2019 by the PMDA. *Axicabtagene ciloleucel:* CD19 directed CAR T-cell product. This product received US FDA approval in October 2017, by EMA in August 2018, and January 2021 by the PMDA. *Lisocabtagene maraleucel:* CD19 directed CAR T-cell product. This product received approval from the US FDA in February 2021, by EMA in April 2022, and March 2021 by the PMDA. *Idecabtagene vicleucel:* BCMA directed CAR T-cell product. This product was developed jointly by two sponsors and received US FDA approval in March 2021, by EMA in August 2021, and Jan 2022 by the PMDA. *Ciltacabtagene autoleucel:* BCMA directed CAR T-cell product. This product received FDA approval in February 2022, by EMA in March 2022, and September 2022 by the PMDA.

Topic	Summary Information
Indication/Disease treated	*Tisagenlecleucel:* Paediatric and young adult patients up to 25 years old with B-cell precursor ALL that is refractory or in second or later relapse. Adult patients with relapsed or refractory large B-cell lymphoma after two or more lines of systemic therapy. Indication for adult patients with relapsed or refractory follicular lymphoma was approved later. *Axicabtagene ciloleucel:* Adult patients with relapsed or refractory large B-cell lymphoma, including DLBCL not otherwise specified, primary mediastinal large B-cell lymphoma, high-grade B-cell lymphoma, and DLBCL arising from follicular lymphoma, after two or more lines of systemic therapy. Indication for adult patients with large B-cell lymphoma that is refractory to first-line chemoimmunotherapy or that relapses within 12 months of first-line chemoimmunotherapy was approved later. *Lisocabtagene maraleucel:* Adult patients with relapsed or refractory large B-cell lymphoma after two or more lines of systemic therapy, including DLBCL not otherwise specified (including DLBCL arising from indolent lymphoma), high-grade B-cell lymphoma, primary mediastinal large B-cell lymphoma, and follicular lymphoma grade 3B. Indication for adult patients with large B-cell lymphoma that is refractory to first-line chemoimmunotherapy or that relapses within 12 months of first-line chemoimmunotherapy was approved later. *Idecabtagene vicleucel:* Adult patients with relapsed or refractory multiple myeloma after four or more prior lines of therapy, including an immunomodulatory agent, a proteasome inhibitor, and an anti-CD38 monoclonal antibody. *Ciltacabtagene autoleucel:* Adult patients with relapsed or refractory multiple myeloma after four or more prior lines of therapy, including a proteasome inhibitor, an immunomodulatory agent, and an anti-CD38 monoclonal antibody.
Where were the study protocols registered?	-
Stage of the medicinal product development lifecycle	Post-authorisation
How did the involvement of RWD/RWE in the study affect the study design at the outset?	Secondary use for PASS by MAH of academia-led patient registries for CAR T-cell therapies.

Topic	Summary Information
What were the data sources used and why were they chosen?	Patient registries of haematopoietic cell transplantation (HCT) and cellular therapy (CT) of the Center for International Blood and Marrow Transplant Research (CIBMTR) in the US, EBMT in Europe, and Japanese Data Center for Hematopoietic Cell Transplantation (JDCHCT) in Japan in collaboration with Japanese Society for Transplantation and Cellular Therapy (JSTCT). (1) The CT registries were developed in addition to HCT registry with long history and high research productivity. Therefore, CT registries of the CIBMTR, EBMT and JSTCT/JDCHCT are expected to have a high capture rate. The registry holders (CIBMTR, EBMT and JSTCT/JDCHCT) are also academic societies representing the therapeutic community. (2, 3) As new genetically modified T-cell (CAR T-cell) therapies were being developed, the three HCT registries of CIBMTR, EBMT and JSTCT/JDCHCT started a project to harmonise survey items for CAR T-cell therapies in 2016. The patient registry's international collaborative activities were effective for CAR T-cell manufacturing company, most of which are global companies. EBMT responded to the EMA's Patient Registry Initiative in 2016; the EBMT registry was qualified as suitable for the collection of data for PASS in 2019.
What were the data analysis methods used? Why were they chosen?	Secondary use of patient registry data, cohort design was used in the studies. (4, 5) The academia-led patient registries for CAR T-cell therapies of the CIBMTR, EBMT and JSTCT/JDCHCT have harmonised survey items and can be a common data collection platform for multiple CAR-T products. The demand for quality assurance in regulatory use is a challenge from the perspective of sustainability, as it can lead to excessive quality requirements and high costs for academic registries.
What legal data protection requirements had to be met in the countries you were working in?	Patient registry protocol, approved by IRB/ethical committee of treating centres. Informed consent forms include not only data being used for research but also individual data to be shared with MAHs and regulatory authorities only for purposes of supporting post-authorisation safety studies. *Legal data protection requirements* US—US FDA guidance on long-term follow up after administration of human gene therapy products outlines the requirements to industry and manufacturers of these products. EU—Good Pharmacovigilance Practice (GVP) module VIII – Post-authorisation safety studies, ENCePP Code of Conduct, General Data Protection Regulation (2016/679), 21 CFR part 11, Electronic records; Electronic signatures, EBMT standard operating procedures. Japan—Ethical Guidelines for Medical and Health Research Involving Human Subjects, Personal Information Protection Law, GPSP Ordinance of Pharmaceuticals and Medical Devices Law, Computerized System Validation Guideline, Electronic records/Electronic signatures Guideline.

Topic	Summary Information
What did you change (if anything) to be in line with ethical considerations?	None
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/RWE study impacts the regulatory/HTA decision?	US: The US FDA regulates all cell and gene therapy products in the US. All products that have genetic manipulation are required to have a long-term follow up plan to assess the development of subsequent neoplasms 15 years after the infusion. The approach was to create a prospective OS that includes a set number of patients and follow them for 15 years. These studies are product- and indication-specific and it is mainly to fulfil a regulatory requirement. The regulatory decision to approve a product takes into consideration the long-term follow up plan. EU: HTA decision processes vary in Europe by country or region. However, the EUnetHTA Joint Action 3 (2016–2021) developed a permanent HTA working structure for Europe, and was succeeded by EUnetHTA 21 (2021–2023) and focused on supporting a future EU HTA system under the HTA regulation. Within this context the EBMT registry underwent a qualification procedure by EUnetHTA, which might in the future assist in providing a more harmonised HTA strategy based on RWD (https://www.eunethta.eu/wp-content/uploads/2021/07/Written-EUnetHTA-recommendations-EBMT-CAR-T_final-report-June-2021_EBMT_Final.pdf). GoCART-coalition (https://thegocartcoalition.com), a multi-stakeholder initiative aims to bring the loose initiatives together. Japan: PASS results by secondary using JSTCT/JDCHCT cellular registry data are planned to be used in regulatory decision making in re-examination system by the PMDA. In Japan the Safety of Regenerative Medicine Act, passed in 2014, delineates a system for conditional, time-limited MA of these products. The MHLW and PMDA have issued a policy regarding a patient registration system for the collection and evaluation of post-marketing information on regenerative medicine products. In 2018, reliability standards for post-authorisation studies were updated to define post-marketing database (revised GPSP ordinance). The appropriate regulatory guidelines for these studies, including guidance for data reliability assurance when performing a database survey or using patient registries for regulatory purposes, have been promulgated.
Conclusion	International collaboration among academia registries led to creating a common data collection platform of RWD for multiple CAR T-cell products. The platform is useful for PASS of products for MAHs, most of which are global companies.

Topic	Summary Information
Published reference	1. Aljurf M, Rizzo JD, Mohty M, et al. Challenges and opportunities for HSCT outcome registries: perspective from international HSCT registries experts. *Bone Marrow Transplant*. 2014;49(8):1016-21. 2. McGrath E. Chabannon C, Terwel S, Bonini C, and Kuball J. Opportunities and challenges associated with the evaluation of chimeric antigen receptor T cells in real-life. *Curr Opin Oncol*. 2020 Sep;32(5):427-433. 3. Atsuta Y, Okamoto S, and Teshima T. Establishment of a Cellular Therapy Registry in Japan, a platform for data sharing for research, industry, and regulatory use. *Japanese Journal of Transplantation and Cellular Therapy*. 2022;11(4):193-198. 4. Pasquini MC, Hu ZH, Curran K, et al. Real-world evidence of tisagenlecleucel for pediatric acute lymphoblastic leukemia and non-Hodgkin lymphoma. *Blood Adv*. 2020 Nov 10;4(21):5414-5424. 5. Jacobson CA, Locke FL, Ma L, et al. Real-World Evidence of Axicabtagene Ciloleucel for the Treatment of Large B Cell Lymphoma in the United States. *Transplant Cell Ther*. 2022 Sep;28(9):581.

H. New rare disease indication using RWE — US FDA Approval Summary: Alpelisib for PIK3CA-related Overgrowth Spectrum (PROS)

Topic	Summary Information
Rationale	This study aimed to assess efficacy and safety of alpelisib, an inhibitor of phosphatidylinositol-3-kinase (PI3K), in the treatment for PIK3CA-related overgrowth spectrum (PROS). PROS is an umbrella term for several ultra-rare clinical entities resulting from somatic activating mutations in the PIK3CA gene, and presents with asymmetric, sporadic overgrowths and vascular malformations that vary in severity from localised overgrowths with minimal morbidity to extensive and potentially life-threatening overgrowths. This case study is based on data from a single-arm retrospective chart review study that enrolled eligible paediatric and adult patients with severe or life-threatening manifestations of PROS who received alpelisib as part of an expanded access programme (EAP) for compassionate use and had available medical chart history.
Study question. What was the research question in the example?	Is the treatment with alpelisib associated with a clinically relevant efficacy signal (i.e. to reduce the volume of up to three target lesions) of at least 20% at week 24 in paediatric and adult patients suffering from severe or life-threatening PROS?
Medicinal product	Alpelisib (BYL719) is an inhibitor of phosphatidylinositol-3-kinase (PI3K) with inhibitory activity predominantly against PI3Kα. Alpelisib in combination with fulvestrant was approved by the US FDA on 24 May 2019 for the treatment of postmenopausal women and men with hormone receptor (HR)-positive, HER2-negative, PIK3CA-mutated, advanced or metastatic breast cancer as detected by an FDA-approved test following progression on or after an endocrine-based regimen. Subsequently, alpelisib in combination with fulvestrant was approved by the EMA on 27 July 2020 for the treatment of postmenopausal women, and men, with HR-positive, HER2-negative, locally advanced or metastatic breast cancer with a PIK3CA mutation after disease progression following endocrine therapy as monotherapy.

Topic	Summary Information
Indication/Disease treated	PIK3CA-related overgrowth spectrum (PROS) is an umbrella term for a group of rare overgrowth disorders that results from somatic gain of function alterations in the PIK3CA gene, activating the phosphatidylinositol/AKT/mTOR pathway. These clinical entities are heterogeneous in both genotype and phenotype but are generally characterised by asymmetric and sporadic lesions that can result in progressive disability. Although the prevalence of PROS is difficult to estimate due to its rarity, recent characterisation and potential misdiagnosis, based on data from NIH Genetics Home Reference and Orphanet, it is estimated to affect about 14 per million patients (i.e. fewer than 5000 patients in the US). Patients have few options for treatment of their disease, which include surgery (debulking or amputation, which are associated with high risk of regrowth), sclerotherapy, endovascular occlusive procedures, off-label use of inhibitors targeting the PI3K/AKT/mTOR pathway and symptomatic treatment.
Where were the study protocols registered?	The study was conducted across seven sites in five countries (Australia, Ireland, France, Spain, and US).
Stage of the medicinal product development lifecycle	The study was conducted after regulatory approvals were received for alpelisib for the treatment of selected patients with advanced or metastatic breast cancer as well as EAP for compassionate use for the treatment of selected patients with PROS.
How did the involvement of RWD/RWE in the study affect the study design at the outset?	This is a retrospective chart review study based only on RWD. To mitigate potential biases, the following features were considered: a) use of a prospectively defined protocol for data collection and statistical analysis plan finalised before the start of data collection, b) the primary endpoint was based on the radiological assessments of a blinded independent central review (BICR), and c) broad eligibility of patients participating in the EAP to reduce selection bias (e.g. all patients who initiated alpelisib before a defined date were considered for enrolment into the study).
What were the data sources used and why were they chosen?	Patient-level data were abstracted from medical charts of all eligible patients who received alpelisib in the context of an EAP for compassionate use at all participating sites. Given the rarity of PROS, its high unmet medical need, and the preliminary promising results published, the data used were derived from an EAP for compassionate use to support the development of alpelisib in this setting in an accelerated fashion. The study was designed with predefined eligibility criteria and included guidance for patient enrolment, data management and monitoring. All sites participating in the EAP were contacted to participate in the study. Data quality assurance and quality control measures (e.g. standardised data abstraction and data validation) were also implemented.

Topic	Summary Information
What were the data analysis methods used? Why were they chosen? What were the advantages? Disadvantages?	The study protocol and statistical analysis plan were developed before data abstraction and therefore the data analysis methods were pre-specified. The primary analysis was descriptive (estimation based). No hypothesis testing was conducted. The primary analysis was performed on all patients with at least one target lesion and an imaging scan performed on the index date (efficacy population) without missing response (complete case analysis). The initial plan was to perform the primary analysis based on efficacy population considering multiple imputation for missing response. However, the MAR assumption required for multiple imputation analysis may not hold and hence as suggested by a health authority, complete case analysis was considered for the primary analysis despite its potential limitations. Sensitivity analyses were performed to investigate the robustness of the results.
What legal data protection requirements had to be met in the countries you were working in?	Data for the study were abstracted from medical charts of patients who were started on alpelisib on or before September 23, 2019, at participating clinical sites in five countries (Australia, France, Ireland, Spain, and US). All patients in the EAP meeting inclusion criteria were eligible to participate in the study. Prior to data abstraction, informed consent was obtained from the participant (or parent/guardian for paediatric patient < 18 years of age), if required by local regulations.
Is the study for internal decision making or part of regulatory/HTA commitment? If the latter, how the RWD/ RWE study impacts the regulatory/HTA decision?	This was used as a pivotal study to support accelerated approval of alpelisib in the US for the treatment of adult and paediatric patients two years of age and older with severe manifestations PROS who require systemic therapy. Substantial evidence of effectiveness for alpelisib in patients with severe manifestations of PROS who require systemic therapy was supported by RWD collected in the study.
Conclusion. Do you have recommendations or key learnings to share?	▶ Early and frequent engagement with Health Authorities is key. ▶ Agreement on endpoints. ▶ Strategy for collection and analyses of RWD to minimise biases and ensure high data quality. ▶ Sampling process of patients. ▶ Agreement on study/studies to confirm clinical benefit. ▶ Transparent and close collaboration with health authorities throughout the review (e.g. sample size, missing data).
Published reference	Canaud G, Lopez Gutierrez JC, Irvine AD, Vabres P, et al. Alpelisib for treatment of patients with *PIK3CA*-related overgrowth spectrum (PROS). *Genet Med*, 2023, 25(12) :100969. https://doi.org/10.1016/j.gim.2023.100969.

Appendix 1 – References

1 World Health Organization COVID-19 dashboard (Website accessed 11 March 2024).

APPENDIX 2.
ICMRA STATEMENT ON INTERNATIONAL COLLABORATION TO ENABLE REAL-WORLD EVIDENCE (RWE) FOR REGULATORY DECISION-MAKING

Background

The role of real-world data (RWD) and real-world evidence (RWE) in supporting the development of medicines across their different stages of development and use is evolving rapidly. However, challenges exist, due for example to heterogeneous data sources, different levels of data quality, and various governance models for data sharing and access. Close collaboration between regulators across the world can help address these challenges. ICMRA can play an important role by catalysing increased cooperation on the use of RWE for regulatory decision-making. The timely work undertaken by regulators and researchers to address the unprecedented challenge of the COVID-19 pandemic, as well as lessons learnt throughout the last two years, have led regulators to establish or reinforce collaborations allowing efficient sharing of data and experience. These collaborations can be further leveraged to medicines regulation beyond the pandemic. In June 2022, EMA, US FDA, and HC co-chaired an ICMRA workshop (programme in Annexe) to share experience on accomplishments and challenges of RWE in medicines regulation, and to identify opportunities for future regulatory collaboration.

Opportunities for collaboration

The June 2022 ICMRA workshop on RWE identified four areas of opportunities for regulator collaboration which could help address common challenges and further enable the integration of RWE into regulatory decision-making.

- Harmonisation of RWD and RWE terminologies:
  - Generate common operational definitions of RWD and RWE, with clear scope and level of granularity (e.g. pertaining to RCTs and OSs);
  - Leverage existing ICH activities, such as M14 on 'General principles on planning and designing pharmacoepidemiological studies that utilise real-world data for safety assessment of a medicine'.
- Convergence on RWD and RWE guidance and best practice, including:
  - Common principles for RWD quality;
  - Metadata to enable characterisation and discoverability of RWD;
  - Suitable scenarios where RWE may contribute to regulatory decision-making, building on existing use-cases;
  - Templates for study protocols/reports that can be used in multiple regulatory jurisdictions.

- Readiness
  - Through the strengthening of international regulatory collaboration on RWE, enable the rapid creation of expert groups on specific topics of interest, including in case of emerging health threats;
  - Foster collaboration on governance and processes to enable the efficient conduct of studies based on RWD from different countries to address important public health challenges.
- Transparency
  - Define common principles and practices for the systematic registration of pre-specified study protocols (including description of feasibility assessments) and study results in publicly available registries;
  - Promote publication of study results in open-source, peer-reviewed journals.

These potential areas for regulatory collaboration on RWD and RWE could be taken forward through a variety of existing fora including ICH, international standardisation bodies, and clusters of interested regulators. ICMRA remains committed to steering this work in the interests of patient health and innovation.

Sessions	Outputs	Chairs and Speakers
RWE terminology	Review of existing definitions of RWD/RWE	**John Concato - US FDA** Andrew Raven - HC
From RWD to RWE	Lessons learnt from RWE evaluations, successes, and pitfalls	**John Concato - US FDA** Gustavo Mendes Lima Santos - ANVISA Boitumelo Semete - SAHPRA Fawaz F. Al-Harbi - SFDA Daniel Lottaz & Lorenzo Hess - Swissmedic
Landscape analysis of international initiatives	Learnings from ICH and other initiatives about challenges and opportunities, gaps, and future activities	**Melissa Kampman - HC** Ron Milo - Weizmann Institute of Science Corinne de Vries - EMA David Moeny - US FDA David Brown - MHRA
Data sources and metadata	Lessons learnt from using different data sources and perspectives for data discoverability (metadata) and data quality assessment	**Xavier Kurz - EMA** Ana Cochino - EMA Sreemanee Dorajoo - HAS Jun Zhao - NMPA/CDE Peter Mol - CBG-MEB
Federated and other Data Networks	Exploration o f existing federated data networks used worldwide including their challenges and opportunities	**Melissa Kampman - HC** Jesper Kjaer - DKMA Azumi Takano - PMDA Patricia Bright - U S FDA
Other topics of interest	Insight into specific topics of interest in the different regions (e.g. pharmacogenomics)	**Catherine Cohet - EMA** Sarah Vaughan - MHRA Maria Gordillo-Maranon - EMA
Conclusion	Draft statement on international coordination of activities to advance RWE	**Peter Arlett - EMA** **Melissa Kampman - HC** **John Concato - US FDA**

APPENDIX 3.
CIOMS WORKING GROUP MEMBERSHIP AND MEETINGS

The CIOMS Working Group XIII on *Real-world data and real-world evidence in regulatory decision making* included the following groups of stakeholders: regulatory authorities, pharmaceutical companies, academics, and HTA bodies.

Regulatory authorities		
Name	Company/Organisation	Country
Alteri, Enrica	Formerly European Medicines Agency (EMA)	Switzerland
Haenisch, Britta	Federal Institute for Drugs and Medical Devices (BfArM)	Germany
Heß, Steffen	Federal Institute for Drugs and Medical Devices (BfArM)	Germany
Irs, Alar	State Agency of Medicines	Estonia
Ishiguro, Akihiro	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Junji, Moriya	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Lambert, Laurie	Canadian Agency for Drugs and Technologies in Health (CADTH)	Canada
Li, Jie	Food and Drug Administration (FDA)	USA
Schiel, Anja	Norwegian Medicines Agency (NoMA) and Scientific Advice Working Party, European Medicines Agency (EMA)	Norway
Soares, Monica	Brazilian Health Regulatory Agency (ANVISA)	Brazil
Wakao, Rika	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Pharmaceutical companies		
Name	Company/Organisation	Country
Aubrun, Elodie	Novartis Pharma AG	Switzerland
Baumfeld Andre, Elodie	Roche	USA
Blackburn, Stella	IQVIA	UK
Boerstoel, Mariette	AstraZeneca	Switzerland
Brookland, Thomas	Roche	USA
Campbell, Ulka	Pfizer	USA
Gomez-Caminero, Andres	Merck Sharp & Dohme	USA
Gomez-Reino, Elisa	Alexion	USA
Iyasu, Solomon	Formerly Merck Sharp & Dohme	USA

Juhaeri, Juhaeri	Sanofi	USA
Machlitt, Andrea	Bayer	USA
Mera, Robertino	Gilead	USA
Rubino, Heather	Pfizer	USA
Wormser, David	Novartis Pharma AG	USA
Zint, Kristina	Boehringer Ingelheim	Germany
Academics and other stakeholders		
Name	Company/Organisation	Country
Atsuta, Yoshiko	Aichi Medical University School of Medicine/Japanese Data Center for Hematopoietic Cell Transplantation	Japan
Azoulay, Laurent	McGill University	USA
Goettsch, Wim	National Health Care Institute, Diemen/Utrecht University	Netherlands
Hennessy, Sean	University of Pennsylvania	USA
Jonsson Funk, Michele	University of North Carolina at Chapel Hill	USA
Mayer, Miguel-Angel	Hospital del Mar Barcelona/Universitat Pompeu Fabra Barcelona	Spain
Nonaka, Takahiro	Osaka Metropolitan University	Japan
Shaw, David	Maastricht University/University of Basel	Switzerland
Stingl, Julia	University Hospital of Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen	Germany
Townend, David	University of London	UK
Wang, Shirley	Harvard Medical School	USA
Working Group members' alternates		
Name	Company/Organisation	Country
Crane, Gracy	Roche	USA
de Luise, Cynthia	Pfizer	USA
Nishioka, Kinue	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Nomura, Manami	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Sato, Daisaku	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
Wicherski, Julia	Federal Institute for Drugs and Medical Devices (BfArM)	Germany
CIOMS		
Name	Company/Organisation	Country
Hill, Sanna	CIOMS	Switzerland
Rägo, Lembit	CIOMS	Switzerland

The Working Group XIII met 19 times from 2020 to 2024. Most of the meetings took place virtually except for the penultimate meeting, which was hybrid in nature although the majority attended in-person in Geneva, Switzerland. The meetings took place as follows:

1. 30 March 2020
2. 29 June 2020
3. 19 August 2020
4. 23-24 September 2020
5. 15 and 17 December 2020
6. 26 February 2021
7. 1 April 2021
8. 6 May 2021
9. 22 June 2021
10. 27 September 2021
11. 3 December 2021
12. 4 February 2022
13. 5 April 2022
14. 7 June 2022
15. 9 September 2022
16. 5 December 2022
17. 30 May 2023
18. 4-5 October 2023
19. 8 April 2024

The CIOMS Working Group XIII Editorial Committee met 34 times from 2022 to 2024, and included the following Working Group XIII members: Yoshiko Atsuta, Sean Hennessy, Juhaeri Juhaeri, Jie Li, and Lembit Rägo, and received support from Sanna Hill.

APPENDIX 4.
LIST OF PUBLIC CONSULTATION COMMENTATORS

Name	Company/Organisation	Country
Aurich, Beate	c4c Consortium	France
Baptiste, Terrell	Gilead Sciences	United States
Baroutsou, Varvara	Ethics Working Group, International Federation of Associations of Pharmaceutical Physicians and Pharmaceutical Medicine	
Bartels, Miranda	Kyowa Kirin Holdings B.V.	the Netherlands
Becker, Sander	Ethics Working Group, International Federation of Associations of Pharmaceutical Physicians and Pharmaceutical Medicine	
Beyrer, Julie	Eli Lilly and Company	
Bivens, Kristin Marie	University of Bern, Switzerland	Switzerland
Bouche, Louise	Sanofi	United Kingdom
Brooke, Nicholas	Patient Focused Medicines Development (PFMD)	Belgium
Brown, Jeffrey	TriNetX and Harvard Medical School, US	United States
Chandler, Rebecca	Coalition for Epidemic Preparedness Innovations (CEPI)	United Kingdom/ Norway
Chandler, Rebecca	ISoP RWE and big data Special Interest Group	
Chhina, Nimi	BioMarin Pharmaceutical	United States
Curry, David	GE2P2 Global Foundation	United States
Dareng, Eileen	AstraZeneca	United Kingdom
de Haart, Karin	IQVIA	the Netherlands
DeTora, Lisa	Hofstra University, US	United States
Edwards, Sarah	University College London, UK	United Kingdom
Ellis, Darcy	Clinical Epidemiology, CSL Behring	Australia
Eltahir, Salma	1. University of Bordeaux, France 2. IQVIA	Sweden
Facey, Karen	FIPRA	United Kingdom
Feng, Sheng	Parexel	United States
Flannery, Brenda	Bristol Myers Squibb	United States
Franken, Andreas	German Medicines Manufacturers' Association (BAH.e.V)	Germany
Fries, Michael	Biostatistics, CSL Behring	Australia
Gillard, Paul	PPD part of Thermo Fisher Scientific	Belgium

Name	Company/Organisation	Country
Gruber, Susan	TL Revolution, Putnam Data Sciences	United States
Haraszkiewicz-Birkemeier, Natalia	BioPharma First Consultancy	the Netherlands
Holbrook, Anne	McMaster University, Canada	Canada
Ida, Fidelia	Clinical Epidemiology, CSL Behring	Australia
Kayode, Gbenga	Clinical Epidemiology, CSL Behring	Australia
Kerpel-Fronius, Sandor	Ethics Working Group, International Federation of Associations of Pharmaceutical Physicians and Pharmaceutical Medicine	
Kurihara, Chieko	Ethics Working Group, International Federation of Associations of Pharmaceutical Physicians and Pharmaceutical Medicine	
Layton, Deborah	AstraZeneca	United Kingdom
Lazdins-Helds, Janis	Independent Adviser	
Lerner-Hillner, Petra	MERCK KGaA	Switzerland
Lumsden, Rebecca	Sanofi	United Kingdom
Maro, Judith	Harvard Pilgrim Health Care Institute and Harvard Medical School, US	United States
Matsuyama, Kotone	Ethics Working Group, International Federation of Associations of Pharmaceutical Physicians and Pharmaceutical Medicine	
McCully, Stuart	Real-World Research Ltd	United Kingdom
Mol, Peter	MetReal cluster, University Medical Center Groningen, University of Groningen, the Netherlands	the Netherlands
Naumann-Winter, Frauke	Federal Institute for Drugs and Medical Devices	Germany
Pai, Sudarshan	Mylan Pharmaceuticals Private Limited, A Viatris Company	India
Pick, Stephanie	German Medicines Manufacturers' Association (BAH.e.V)	Germany
Polidori, Piera	European Association of Hospital Pharmacists (EAHP)	Belgium
Redman, Barbara	GE2P2 Global Foundation	United States
Robins, Deborah	Independent consumer volunteer/advocate	Australia
Saio, Takeo	Fuji Toranomon Orthopedic Hospital	Japan
Schneider, Muriel	Global Self-Care Federation	Switzerland
Tan-Koi, Wei Chuen	Centre of Regulatory Excellence, Duke-NUS	Singapore
Thakur, Rohit	Novo Nordisk A/S	India
Toh, Darren	Harvard Medical School, US	United States
van der Laan, Mark	TL Revolution, UC Berkeley	United States
van der Laan, Mark	University of California, Berkeley, US	United States

Name	Company/Organisation	Country
Wilson, Andy	Parexel	United States
Younes, Fariha	Sanofi	United Kingdom
Young, Michael	Massachusetts General Hospital and Harvard Medical School	United States
Zint, Kristina	Boehringer Ingelheim International GmbH	Germany
Collated comments	CADTH	Canada
Collated comments	CADTH-Health Canada RWE Steering Committee	Canada
Collated comments	European Federation of Pharmaceutical Industries and Associations (EFPIA)	
Collated comments	International Society for Pharmacoepidemiology (ISPE)	
Collated comments	International Society for Pharmacoepidemiology (ISPE) Collaborative Special Interest Group	
Collated comments	US Food and Drug Administration	United States

In recent years, many medicines regulatory agencies have expressed increased willingness to consider real-world evidence (RWE), that derives from the review and/or analysis of real-world data (RWD), to support claims of efficacy or effectiveness as well as of safety. This increased willingness is changing the regulatory environment in which RWE is generated and used. This consensus report aims to describe the potential use of RWE for decision making; RWD and data sources; key scientific considerations in the generation of RWE; and ethical and legal issues in using RWD.

The intended audience for this report includes medicinal product regulators, healthcare payers, health care and medicinal products industries, researchers, bioethicists, patients and health care professionals. This report was developed to inform discussions about the use of RWD and RWE for regulatory and health care decision making, including decisions to make a product available for use (authorisation), to cover the costs of its use (reimbursement), and to use a product for a particular patient (clinical use).

This report reflects the opinions of the Council for International Organizations of Medical Sciences (CIOMS) Working Group XIII on *Real-world data and real-world evidence in regulatory decision making*, and it was finalised after considering comments received during a public consultation.

Real-world data and real-world evidence in regulatory decision making
Report of the CIOMS Working Group XIII. Geneva, Switzerland.
Council for International Organizations of Medical Sciences (CIOMS), 2024.

CIOMS publications may be obtained directly from CIOMS through its publications e-module at https://cioms.ch/publications/. Further information can be obtained from CIOMS, P.O. Box 2100, CH-1211 Geneva 2, Switzerland, www.cioms.ch, e-mail: info@cioms.ch.

CIOMS
1949-2024
COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES

ISBN 978-92-9036-106-0