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

CIOMS Working Group report
Draft, 6 June 2026

This report was posted for comment on 6 June 2023 at: https://cioms.ch/working-groups/real-world-data-and-real-world-evidence-in-regulatory-decision-making/.

The CIOMS Working Group (WG) XIII welcomes your input to the report, or any parts of it. A list of WG XIII members can be found on the CIOMS website. A detailed list will be appended to the final report.

Please note that the layout will be improved in the final version, and best efforts will be made to correct remaining typographical and/or grammatical errors, as well those pertaining to references.

Permissions are being sought to reproduce some of the illustrative materials included in this report. We welcome responses from organisations that own any of these materials and have not yet been contacted in this regard.

Please submit your comments using the form posted on the CIOMS website at https://cioms.ch/working-groups/real-world-data-and-real-world-evidence-in-regulatory-decision-making/.
The timeline for submission of comments is 14 July 2023.

Thank you.

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), 2023.

DOI *number tbc*

This publication is freely available on the CIOMS website at: https://cioms.ch/working-groups/real-world-data-and-real-world-evidence-in-regulatory-decision-making/

Acknowledgements

(to follow)

Table of contents

Chapter 1: Uses of real-world evidence for decision making during the product lifecycle 22

1.2 Evidentiary requirements by regulators or HTAs 24

1.3 Planning for RWE in each phase of product development 26

1.4 RWE use in lifecycle of the development of medical products 28

1.5 Adapt good clinical practices concepts of data integrity to RWD 33

1.6 Evidence generation presentation and communication 33

1.7 Engaging with regulators 34

Chapter 2: Real-world data sources 43

2.1 Traditional sources 43

2.2 Emerging data sources 48

Chapter 3: Key scientific considerations in regulatory real-world evidence generation 54

3.1 Data source and data quality, integrity, transparency for data transformations, fitness for purpose 54

3.2 Study design and methods 56

3.3 Considerations for statistical analysis in a RWD setting 69

3.4 Evidence-generation process, study registration, transparent reporting, audit trails and responsible communication 73

3.5 Reproducibility of RWD studies 76

3.6 Agreement between multiple RWD studies and RCTs 77

3.7 Quality of RWD studies 77

Chapter 4: Ethical and legal issues in using RWD 83

4.1 The current normative landscape 83

4.2 Ethical arguments for incorporating more RWE 85

4.3 Potential ethical issues in using RWD 87

4.4 RWD, privacy and data protection 87

4.5 Summary 93

Chapter 5 Conclusions and future directions 97

APPENDIX 1: Case studies 100

A. Fosdenopterin approved for treatment of a rare, genetic disease with external control data from a natural history disease study 100

B. Comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with different COVID-19 vaccines in an international network cohort study 102

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 105

D. N-Nitrosodimethylamine (NDMA)-contaminated valsartan and the risk of cancer 108

E. Cardiovascular risk of urate-lowering drugs: a study using the National Database of Health Insurance Claims and Specific Health Check-ups of Japan 110

F. Nested case-control study utilising MID-NET® on thrombocytopenia associated with pegfilgrastim in patients treated with antineoplastic agents 112

APPENDIX 2: ICMRA statement on international collaboration to enable real-world evidence (RWE) for regulatory decision-making 114

APPENDIX 3: CIOMS Working Group membership and meetings 116

APPENDIX 4: List of commentators (following public consultation) 117

List of boxes, figures and tables

Box 1: ISPE/ISPOR taskforce recommendations for HETE

Figure 1: Potential use of RWE in each core regulatory review process

Figure 2: Process proposed by TRUST4RD Tool

Figure 3: Examples of using real-world data (RWD) in the stages of the drug lifecycle

Figure 4: Reliance on RWD in representative types of study design

Table 1: Some definitions of real-world data

1

Abbreviations

AESI	Adverse Events of Special Interest
ANVISA	Brazilian Health Regulatory Agency
BLA	Biologics License Application
BMI	Body mass index
CADTH	Canadian Agency for Drugs and Technologies in Health
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
eCRF	Electronic case report form
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
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).
HCP	Health care professional *or* health care provider
HETE	Hypothesis evaluating treatment effectiveness
HMA	Heads of Medicines Agencies

HTA	Health technology assessment
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICMRA	International Coalition of Medicines Regulatory Authorities
IRB	Institutional Review Board
ISPE	International Society for Pharmacoepidemiology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
IVD	In-vitro diagnostic medical device
MA	Marketing Authorisation
MAH	Marketing Authorisation Holder
MAR	Missing at Random
MCAR	Missing Completely at Random
MHLW	Ministry of Health, Labour and Welfare (of Japan)
MIHARI	Medical Information for Risk Assessment Initiative
MNAR	Missing not at Random
MoCD	Molybdenum cofactor deficiency
NDA	New Drug Application
NDMA	N-nitrosodimethylamine
NICE	National Institute for Health and Care Excellence
NMPA	National Medical Products Administration (of People's Republic of China)
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
PIPEDA	Personal Information Protection and Electronic Documents Act
PMDA	Pharmaceuticals and Medical Devices Agency (of Japan)
PRO	Patient-reported outcome

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
TGA	Therapeutic Goods Administration (of Australia)
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TRUST4RD	Tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases
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

Since the early 1960s, national drug regulatory agencies have required adequate and well-controlled clinical studies as evidence of efficacy as a precondition for approving a new medicinal product. This evidence has usually taken the form of the results of randomised controlled trials (RCTs) that compare the new treatment to an inactive placebo. As RCTs are designed to demonstrate efficacy, they are often too small to assure safety with respect to rare adverse effects. Recognising this, regulatory agencies have for many years accepted what we now call real-world evidence (RWE) derived from data collected outside of RCTs to fulfil post-approval safety requirements.

Despite the many strengths of RCTs, trialists often have difficulty achieving enrolment goals, particularly when evaluating treatments for rare conditions. Further, the highly controlled conditions in which many pre-approval RCTs are performed can limit their generalisability. Responding to these challenges, many drug regulatory agencies have in recent years expressed willingness to consider RWE to support claims of efficacy as well as safety. This willingness is producing rapid changes in 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 who are 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 uses of RWE for decision making during the product lifecycle (Chapter 1), describes RWD and its sources (Chapter 2), discusses key scientific considerations in the generation of RWE for regulatory use (Chapter 3), discusses ethical and legal issues in the generation of RWD (Chapter 4), and provides a summary 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

Decisions about the availability, coverage, and use of medical treatments decisions are, or should be based on evidence, and this evidence must be evaluated and weighed by various actors at different stages.

For example, regulators decide whether a medicinal product should be authorised for use, in which conditions or therapeutic indications, and for which patients. Healthcare payers decide whether an authorised medicinal product should be covered, specifically for which medical condition, and at what price. Health care providers (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. All of these decisions rely on evidence about the product's benefits and risks. To allow for the most informed decisions, this evidence needs to be valid and unbiased, or if it is 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) were 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. However, a limitation of typical pre-approval clinical trials is that historically they have tended to enrol subjects who were not always representative of the population who will use the product once it is approved. This has raised continuing questions about whether the resultant findings are generalisable to the sorts of patients, clinicians, and situations that are more commonly seen in the real world. With the evolution of availability and accessibility of real-world data (RWD) as well as evolving methods for the design and analysis of non-randomised studies, the role of RWD in clinical development and informing regulatory, coverage, and utilisation decisions has increased in recent years.

The Council for International Organisations of Medical Sciences (CIOMS) has developed this consensus report to inform discussions about the use of RWD and Real World Evidence (RWE) for regulatory and healthcare 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).

We propose to define RWD as **health-related data collected from patients or caregivers in routine clinical practice without a study-determined intervention. RWD can come from a wide variety of sources such as healthcare claims and health records, registries, patient reported outcomes, digital tools/wearables/mobile devices. Data collected can include clinical and economic outcomes, patient-reported outcomes, such as disease activity and quality of life, and resource utilisation.**

RWE is evidence derived from the review and/or analysis of RWD.

The intended audience for this report includes medicinal product regulators, healthcare payers, healthcare and medicinal products industries, researchers, bioethicists, patients and HCPs, who produce RWE or use it to inform regulatory, reimbursement, or clinical decisions. This guidance 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 guidance is the use of RWE to evaluate medicinal products, i.e. drugs and biologicals, many of the considerations discussed in this guidance can be applied to medical devices as well.

A variety of stakeholders are 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, and decline.

This report covers the following areas:

- Regulatory potential of RWE and current controversies and challenges;
- Uses of RWE for decision making during the product lifecycle;
- RWD and data sources;
- Key scientific considerations in regulatory RWE generation;
- Ethical and legal issues in using RWD.

Several stakeholders make decisions along a medical product's lifecycle:

The concept of benefit-risk assessment is used by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA), as well as other regulatory agencies, in order to make approval decisions. The structured benefit-risk assessment is also mentioned in International Council for Harmonisation (ICH) guidance, 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 evidence at this stage is usually derived from RCTs.

In a health technology assessment (HTA), the intended and unintended consequences of using a new health technology compared to existing alternatives may be examined. Ascertainment of value is generally based on an integration of 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.

In healthcare, a payer is a person, organisation, or entity that pays for the care services administered 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 the ascertainment of value. Globally, the role of payers is to determine the access of drugs based on reimbursement, budget and pricing.

Patients and providers of care 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 the 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.

Marketing authorisation holders provide evidence to answer questions posed by other stakeholders. This data can come from a variety of sources including RWD.

In general, the totality of 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, as regulators, HTA organisations and payers in different jurisdictions may have different opinions on the value of RWD/RWE.

Regulators are continuously working on providing requirements and recommendations to improve and structure the use of RWD in decision making.

A strategy for addressing evidence gaps should cover all types of evidence generation, including randomised trials and non-randomised studies, and should be based on the research question of interest motivated by the evidence needed by different stakeholders. Common stakeholder requirements/expectations are high-quality data/information and reliability, access and being able to understand the information.

For RWE to influence or support regulatory decision making, all stakeholders, including sponsors, regulators, and HTAs need to implement a transparent process of planning, assessing and reporting of RWE. Transparency of the research processes is key to enabling 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.

Historically, health care databases have been used mainly 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 evaluation of potential safety issues identified in other data sources (signal confirmation or refinement).

In the setting of traditional clinical trials or observational studies which collect data according to the research plan, the data collection phase is included in the research. Thus, data items to be collected and their definitions are designed prior to data collection.

In contrast, in RWE generation/secondary use of existing data or existing database/platform is common. RWD data sources are often created for different purposes, for example to collect data for healthcare or administrative purposes, and the majority of them have not considered research uses at the development of the database. This means that they may or may not be fit for research purposes.

Especially in the secondary use of existing data, it is critical whether the key variables (exposure, outcomes/endpoints, demographic characteristics, and potential confounders) required to answer the clinical questions of the study are reliably collected in the selected data source. If the required variables are not reliably collected in the data source, one could investigate for the possibility of additional data collection.

A strong argument can be made to expand the use of RWD/RWE for the assessment of product effectiveness to support regulatory decisions versus only relying on RCTs. One can assess product effectiveness in a much broader and diverse patient population that reflects settings and patients who will use the product post-approval (e.g. broader range in age, race/ethnicity, comorbidity, disease severity, concomitant medications). One can study a much larger number of patients and for longer durations to increase the potential to detect rare safety outcomes, drug-drug interactions and longer-term effectiveness and safety outcomes. Finally, RWD/RWE studies are less resource intensive as compared to RCTs.

Before considering whether or not to use RWD in a study to support regulatory approval, it is imperative to start with the determination of the research question and the clinical context for the decision. Once these two pieces of information are clarified, one can begin to determine the critical data elements that are needed, evaluate possible data sources that enable the accurate assessment of the eligible target study population, treatment exposures, relevant clinical outcomes, covariates and appropriate study design choice.

Regulatory decisions affecting public health in the form of marketing authorisation approvals and to some extent also reimbursement decisions, have traditionally been based on RCTs for which rigorous criteria to ensure data integrity have been developed. This includes, for example registration of protocols, pre-specifying analysis, blinding subjects, investigators, endpoint adjudicators and analysts, as well as publication and results disclosure.

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

Recent regulatory approvals based on RWE have 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.

In the perspective of a wider use of RWD, leading to its own important contribution to regulatory decision making, one must also consider ethical implications.

Ethics concern what one ought to do at a deeper level than simply because the rule requires it, even by the consensus of democratic opinion. Ethics makes a fundamental appeal to the *rightness* of an action that transcends the particulars of the rule.

The move toward broader use of RWE to evaluate efficacy as well as safety is justified not only by a need for stronger 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 regarding efficacy from RCTs may not translate into evidence regarding effectiveness in clinical care. This is because the actual patient population is often not well represented by typical participants in RCTs, who are often younger and healthier than many patient groups treated in daily practice. RCTs also tend to under-report harm, further weakening the evidence base for real-world clinical care.

RWE is increasingly used in practice, and this often takes place without any ethical or legal framework specific to use of RWD being in place, even if frameworks for clinical trials exist in all jurisdictions. Particularly in the context of the COVID-19 pandemic, personal data was used to inform decision making on a scale not seen before.

With the exception of 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. 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. However, any such change in paradigm cannot be accomplished by diktat; societal discussion would have to precede any such legislative change. As a starting point it is worth considering whether informed consent is necessary as an ethics standard in data use in research. Clearly, RCTs work with this standard. Many RWE and RWD do not work easily with a presumption of informed consent, as they depend on large, secondary use of already gathered data. Is informed consent necessary in all ethics theories?

On the other hand, highly regulated areas such as medical research, with multiple safeguards and independent scrutiny are made challenging to negotiate and undertake. RWD is in danger of being so restricted by data protection law that medical research becomes impossible, 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 knowledge that the RWD research pursues are clearly in the public interest and in the interests of protecting human dignity.

When evaluating the safety of drugs on the market, public interest conceptually trumps individual privacy claims. Most legislation that regulates situations where personal data are processed (for example, the Clinical Trials Regulation in the EU) defer to the General Data Protection Regulation (GDPR) to govern the processing of personal data. When it comes to medicines' safety, it is possible to overrule the general data protection regime. This makes for an interesting anomaly in RWD processing - that processing for effectiveness research must be GDPR compliant, whereas processing in relation to safety questions can be undertaken in some jurisdictions with a rather different approach. Thus, individual agency can be overridden for solidarity needs where there is a political will. One could argue that there is an overriding public interest in establishing not only the safety, but also the effectiveness of a product on the basis of RWD.

RCT evidence is still important, but its focus on "perfect" patients who are often highly unrepresentative of the populations in whom new drugs and other interventions will be used, combined with almost complete neglect of some underserved populations such as pregnant women, older patients and minoritised ethnic groups, and the specific issue of the efficacy-effectiveness gap, mean that using RWE to augment RCT evidence is an ethical imperative.

Many treatments are currently prescribed based on old and unrepresentative RCT evidence. As a result, patients may be prescribed drugs that will not help them, or at least will not help them as much as they and the HCP think. Further, such medicines may cause more harm than anticipated. This means that the principles of beneficence and nonmaleficence are both threatened by failure to use RWD. In turn, it means that if HCPs and patients do not know this, then decisions made may be uninformed, threatening individual autonomy. At a larger scale, use of unrepresentative data across health systems threatens the principle of justice by distributing resources according to similarly flawed decisions. Equally, of course, any 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.

In turn, if it is vital to use RWE more broadly, ethical frameworks, guidance, regulations and legislation must be future-proofed to enable the use of 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.

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 via ensuring informed decision making about medicines, rather than traditional protection of autonomy by keeping personal data siloed and sealed off, are likely to be paramount in increasing utilisation of RWE in an ethically robust manner.

To answer the regulatory, normative and governance questions posed by RWD, we cannot rely on the current political approach that avoids hard moral questions. Only by opening the debate to explore the competing interests of all stakeholders and respecting the concerns and hopes of all parties, at an international level and without any prejudice in favour of the economically rich countries and individuals, can the environment that RWD requires be created. Ironically, the solution is available in plain sight in the current legislation - it is within our grasp. What seems beyond our reach is the will to ask the most important ethical questions. What responsibility do I have to others? What responsibility do I have to produce robust, honest science? What responsibility do I have to ensure access to healthcare products as a part of a right to healthcare? What is my commercial responsibility in that regard? What is my responsibility as a patient and as a member of the public in that regard? What duty of confidence do I owe to anyone whose data I process? What can I demand about “my data”? Can I really demand absolute privacy? Confidentiality conceptually offers the negotiated terms by which information can be used for specified purposes. The purpose of data protection legislation is not to shut down or prohibit the processing of personal data, but rather to regulate it in such a way as to create an appropriate balance of safeguards for the processing of personal data for different, legitimate ends. In that respect, it is probably more appropriate to use the term “confidentiality” when discussing use of personal data for research purposes. The term has strong links to professional duty - to the duty to place one’s clients’ interests before one’s own in acting in a professional capacity. A shift away from a privacy debate to a confidentiality debate offers an opportunity to re-focus the discussion, back to the starting point of asking how to enable data to be processed for legitimate ends and how to safeguard legitimate interests. The professionalisation of researchers, as is perhaps emerging in the drive to address research integrity, cannot come too soon to assist in this re-evaluation of what data protection is seeking to achieve, particularly in terms of using RWD.

This report indicates that it is possible, and indeed necessary, to expand the use of RWD/RWE for regulatory decision making all along the medical product’s lifecycle. It describes the needs of the different stakeholders along the process; it discusses the available data sources, their foreseeable development, their strengths and limitations; it examines what are the key scientific considerations to make in planning RWE generation; and last but not least, it presents ethical and legal perspectives in RWE generation and utilisation.

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] . For example, regulators decide whether a medicinal product should be authorised for use, in which conditions or therapeutic indications, and for which patients. Healthcare payers decide whether an authorised medicinal product should be reimbursed, to whom, and at what price. 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) were 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. Subjects 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 limitation of typical pre-approval clinical trials is that historically they have tended to enrol subjects who were not always representative of the population who will 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 drug or intervention is used in real-world circumstances (effectiveness).[3] While 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 actual target patient population, most pre-approval trials are still performed in relatively selected patient populations, who are treated by highly selected clinical investigators. This has raised continuing questions about whether the resultant findings are generalisable to the sorts of patients, clinicians, and situations that are more commonly seen in the real world. Another limitation is that RCTs are designed with sufficient sample size to assess efficacy, and thus may not have enough statistical power to detect uncommon safety issues. To detect such safety issues in the real-world setting and address them appropriately, studies utilising real-world data (RWD) are needed.

The two limitations mentioned above show how studies using RWD are necessary to address issues for which RCTs are not suitable. It is also important to note that, regardless of RCTs' limitations, studies using RWD, if designed properly and analysed using appropriate methods, can also generate valid evidence, provided that certain assumptions (e.g. no unmeasured confounding) are met.

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: [4]

Organisation	Definition of real-world data
International Society for Pharmacoeconomics and Outcomes Research (ISPOR)[5]	RWD are the 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[6]	…data collected during the routine delivery of care and its reimbursement. This type of data, referred to as real-world data, includes patient registries, EHRs, healthcare claims databases, and patient-generated data and is defined by its production outside of a research study
Innovative Medicines Initiative Get Real Project[7]	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.

Many definitions of RWD are narrow and binary, referring to health care data used for decision making that are not collected in conventional RCTs. Others define RWD more broadly as data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.[8]

We propose to define RWD as **health-related data collected from patients or caregivers in routine clinical practice without a study-determined intervention. RWD can come from a wide variety of sources such as healthcare claims and health records, registries, patient reported outcomes, digital tools/wearables/mobile devices. Data collected can include clinical and economic outcomes, patient-reported outcomes (PROs), such as disease activity and quality of life (QoL), and resource utilisation.**

RWE is evidence derived from the review and/or analysis of RWD.[9]

One 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 to be unethical occurred with avelumab for the treatment of

Merkel Cell carcinoma. At the time 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.[10] Other reasons why a RCT may not be feasible is that patients 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. For all these reasons, some have argued that decision makers should be more flexible in what evidence they accept, 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 study 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 possible study designs to answer a particular question now covers a wide spectrum of possibilities. By definition, 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 manners. However, it is important to note that even investigational products can generate RWD (for example in the frame of compassionate use programmes), and their development can be complemented 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 required the US Food and Drug Administration (FDA) to consider the potential for RWE to evaluate extensions of an existing indication, but not for initial indications.[11] Given that it is generally much less expensive to develop RWE than to perform RCTs to evaluate efficacy, the medicinal products industry has a significant financial incentive to use RWE to support new product indications. Further, the use of RWE to support initial marketing authorisations (MAs) has been tentatively introduced, most frequently in the context of a single-arm trial with a "synthetic control arm", consisting of simulated patients or patients from outside of 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 authors have argued 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,"[12] even though approval decisions by regulatory agencies (including the US FDA[13]) have sometimes been based on non-randomised evidence even before the 21st Century Cures Act was passed. The Council for International Organisations of Medical Sciences (CIOMS) has developed this consensus report to inform discussions about the use of RWD and RWE for regulatory and healthcare 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).

Using RWD is strongly justified on ethical grounds because relying entirely on RCT data could undermine patient autonomy and cause harm, However, its use raises ethical and legal issues which are also addressed in Chapter 4 of this report. The primary issues are patient consent to the use of their data, privacy and data protection.

Regulatory potential of RWE and current controversies and challenges

RCTs have long been recognised as the mainstay for evaluating the efficacy of a medicinal product and are often a prerequisite for obtaining a licence to market a medicine in regulated countries. Randomisation reduces the possibility of imbalances among treatment groups, which can lead to biased study results. The inclusion and exclusion criteria of the RCTs are often relaxed as the investigational product progresses along its development pathway. However, with few exceptions, e.g. vaccines or preventive treatments, the patients who enrol in pre-approval clinical trials are not representative of those seen in a typical doctor's surgery or office. As Eichler et al 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 data from observational studies."[14]

The uncertainties that Eichler et al refer to concern the potential benefits and risks, as well as how a medicine will perform and be utilised 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.[15] In contrast, 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 women, 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 post 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 post-approval, and explain the measures that are taken in order to prevent or minimise the medicine's risks in patients. RMPs may also include mandated studies on aspects of efficacy.[16]

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, the US FDA has created 'The Framework for FDA's Real-World Evidence Program'[17], which clarifies how the agency evaluates adequacy and applicability of types of RWD and RWE for their regulatory decision making.

Similarly, in 2017 the 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 identify how to optimally utilise big data in support of innovation and public health in the European Union (EU).[18]

In addition, in July 2020, EMA issued for consultation their Guideline on registry-based studies. It focuses on studies using registries as a data source with a possible regulatory purpose.

Typically, RWE has been used to fulfil post-approval requirements and conduct long-term follow-up studies if there is remaining uncertainty about risks at the time of approval. Increasingly though, RWD/RWE is applied to capture clinical outcomes in pragmatic and large simple trials. More and more it is also used to provide natural history of disease information to be used as external controls in situations where the use of a randomised comparator arm is impractical or unethical, such as oncology 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 demonstrating 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, and 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 not always allow 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, concerns that 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. Another important factor hindering adoption is that despite the efforts mentioned above, no consistent standards or guidelines exist on how to apply and weigh the RWE in regulatory submissions.

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 is not a large enough patient base, or enough time to gather evidence for approval considerations the traditional way. In such circumstances, RWE can inform about the benefit-risk balance in the target population.

With the 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

CIOMS is an international, non-governmental, non-profit organisation established jointly by the World Health Organisation (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. CIOMS' mission is to advance public health through guidance on health research and policy including ethics, medicinal product development and safety.

The intended audience for this report includes medicinal product regulators, healthcare payers, healthcare and medicinal products industries, researchers, bioethicists, patients and HCPs, who produce RWE or use it to inform regulatory, reimbursement, or clinical decisions. This guidance 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 guidance is use of RWE to evaluate medicinal products, many of the considerations discussed in this guidance can also be applied to medical devices, as well.

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 introductory chapter:

582
583 - Chapter 1 addresses uses of RWE during the product lifecycle;
584 - Chapter 2 addresses RWD and data sources;
585 - Chapter 3 discusses key scientific considerations in regulatory RWE generation;
586 - Chapter 4 addresses ethics, governance and related issues;
587 - Chapter 5 provides conclusion and future directions.

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] *Goettsch W, Makady A. The GetReal Institute. Glossary of Definitions of Common Terms. (PDF accessed 15 September 2022).*

[5] *The Professional Society for Health Economics and Outcomes Research (ISPOR). About Real-World Evidence. (Website accessed 15 September 2022).*

[6] *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. https://doi.org/10.1016/j.jval.2021.07.003*

[7] *Garrison Jr LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR real-world data task force report. Value in Health. 2007;10:5:326-335. https://doi.org/10.1111/j.1524-4733.2007.00186.x*

[8] *The United States Food and Drug Administration (FDA). Real-World Evidence. (Website accessed 15 September 2022).*

[9] *Goettsch W, Makady A. IMI-GetReal Consortium. Glossary of definitions of common terms. (PDF accessed 15 September 2022).*

[10] *European Medicines Agency (EMA): Assessment Report. Bavencio. EMA/496529/2017 (PDF accessed 15 September 2022).*

[11] *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).*

[12] *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*

[13] *Belson NA. FDA's Historical Use of "Real World Evidence." Food and Drug Law Institute. (Website accessed 15 October 2022).*

[14] *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*

[15] *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*

[16] *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).*

[17] *The United States Food and Drug Administration (FDA). Framework for FDA's Real-World Evidence Program. (PDF accessed 2 May 2023).*

[18] *European Medicines Agency (EMA). How we work: Big Data. (Website accessed 15 October 2022).*

588

Chapter 1: Uses of real-world evidence for decision making during the product lifecycle

Health-related RWE, which can be derived from RWD, have the potential to 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[19]) 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 utilise 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.

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

Next, we will discuss frameworks that may be used for acceptance of RWE by stakeholders, including examples of specific frameworks from individual countries and how such frameworks can adapt to be responsive 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.

1.1.1 Regulators, HTA and payers: variety of stakeholders

A variety of stakeholders are 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, and decline. Moreover, within any given health system, they may have divergent views on the potential role of RWE in informing decision-making[20]. In this chapter, roles in relation to RWE and decision making are considered for the following types of stakeholders: regulators, HTA bodies, payers, clinicians, patients and pharmaceutical companies.

Regulators

The role of regulatory bodies, such as the EMA and US FDA, is to authorise entry of a drug into their respective market based on the determination of a positive benefit-risk balance for a specific indication. The fundamental goal of structured benefit-risk assessment is to ensure that the benefits of the drug outweigh the risks throughout its lifecycle. Continual assessment and monitoring of the benefit-risk balance necessitates the ability to evaluate different types of data from multiple sources (see section 1.6.2 on Transparency and disclosure of protocol on RWE reporting below). The concept of benefit-risk assessment is used by the US FDA and EMA, as well as other regulatory agencies. In fact, the effects table, which gives a summary of the favourable and unfavourable effects measured for the alternative(s) and comparator(s) that were taken into account by the regulators, along with descriptions of their uncertainties, and is used by US

FDA and EMA for the structured benefit-risk assessment, has been a standard part of the review. The structured benefit-risk assessment is also mentioned in International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidance, 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.

Health Technology Assessment organisations

CIOMS Working Group XIII defines HTA as "a multidisciplinary 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, program 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 drug 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.[21] In a health technology assessment (HTA), the intended and unintended consequences of using a new health technology compared to existing alternatives may be examined. 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;[22] 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.[23] Ascertainment of value is generally based on an integration of 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.

Payers

In healthcare, a payer is a person, organisation, or entity that pays for the care services administered 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 the ascertainment of value. Globally, the role of payers is to determine the access of drugs based on reimbursement, budget and pricing. Depending on the local established healthcare system, different models exist such as single payer (e.g. as seen in Canada, the UK, or Taiwan) or hybrid models (e.g. as seen in 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 drug 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 can lack 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.

Patients and physicians

The ultimate stakeholders are, of course, patients and their physicians who consume and prescribe these medicines to hopefully improve health and wellbeing. The ultimate goal of informed decision making is to promote treatments to individuals that benefit the most and in the safest possible manner. Patients and providers of care 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 the 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.[24,25]

Pharmaceutical companies, MAHs, and other product developers

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. 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 or HTAs

1.2.1 Frameworks and guidances for RWE by the regulators

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/RWE in the decision-making process depending on the context. In general, the totality of 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 as regulators, HTA organisations and payers in different jurisdictions may have different opinions on the value of RWD/RWE.

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. In December 2018, the US FDA published a Framework[26] for evaluating the potential use of RWE to help support the approval of a new indication for a drug already approved or to help support or satisfy drug 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 conducted 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 considerations of use of electronic health records and claims databases, including recommendations on how to select appropriate RWD sources and to define and validate study variables[27]
- "Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products" is a guidance for use of 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[28]

- "Data Standards for Drug and Biological Product Submissions Containing Real-World Data" focuses on US FDA-supported data standards in drug submissions with data derived from RWD to promote compliance with relevant legal requirements[29]
- "Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drugs and Biological Products" provides US FDA's current thinking regarding regulatory considerations for non-interventional studies involving the use of RWD[30]

In September 2022, the US FDA published the final guidance of "Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products: Guidance for Industry."[31] In early 2023, a draft guidance was published on externally controlled trials.[32]

In Europe, post-authorisation efficacy studies (PAES) are in some instances requested by EMA to generate evidence needed for standard benefit-risk assessment, or at least complementing 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 post-authorisation. EMA has developed an associated scientific guidance to support MAH in the design of PAES.[33]

In 2015, the 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 a guidance on the use of registry-based studies.[34]

Opportunities for improvement in 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.[35] Big Data as discussed by the taskforce included RWD such as 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 taskforce concluded with 10 priority recommendations[36] 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 Task Force report. In the current context of lack of specific guidance for the use of RWD and RWE in pre-approval setting EMA encourages the 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 utilising RWD in clinical trials (ICH E8 R1 and E6 R3) and, the guideline on general principles on pharmacoepidemiological studies that utilise RWD for safety assessment of medicines (ICH M14). However, there seems no overarching ICH guideline that refers to the various guidances that explain how RWD can be used to support clinical trials designs and drug development.[37]

1.2.2 Considerations by HTAs

In the context of HTA and decisions concerning reimbursement, data derived from real-world sources have been used to contextualise information to a specific regional healthcare setting, but initiatives to generate RWE to fill gaps in evidence are increasing.[38] For example, the Commissioning through Evaluation program in England enables new clinical and patient

experience data 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 drugs treating conditions of high and unmet clinical need. Different regions around the world such as Asia, Canada, and the UK are developing and publishing their own frameworks to guide the use, generation, reporting and appraisal of RWE for decision making.[39,40] In 2022, NICE published its real-world evidence framework.[41] 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 used to collect and analyse the data was not considered appropriate or the quality of the data not of an acceptable level.

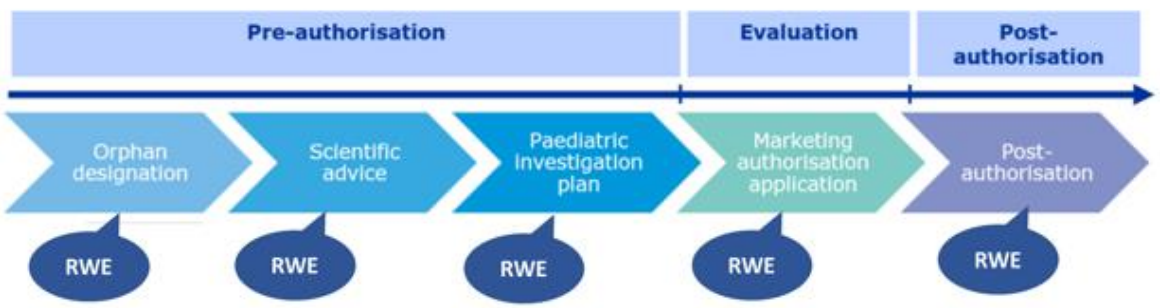
For HTA organisations, local and regional differences in approaches to drug value assessment present additional complexity for drug manufacturers and developers. In the current environment, it is almost impossible for sponsors involved in new product commercialisation to have a common global evidence 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 cancer drug appraisals by UK National Institute for Health and Care Excellence (NICE) from 2011 to 2018,[42] 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 RWE in each phase of product development

1.3.1 Relevant decision points in product lifecycle

Ideally, for each development program, the evidence needed for regulatory approval, including RWE, should be established by the sponsor 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 RWE in each core regulatory review process
Source: Modified from an original EMA figure.[43]

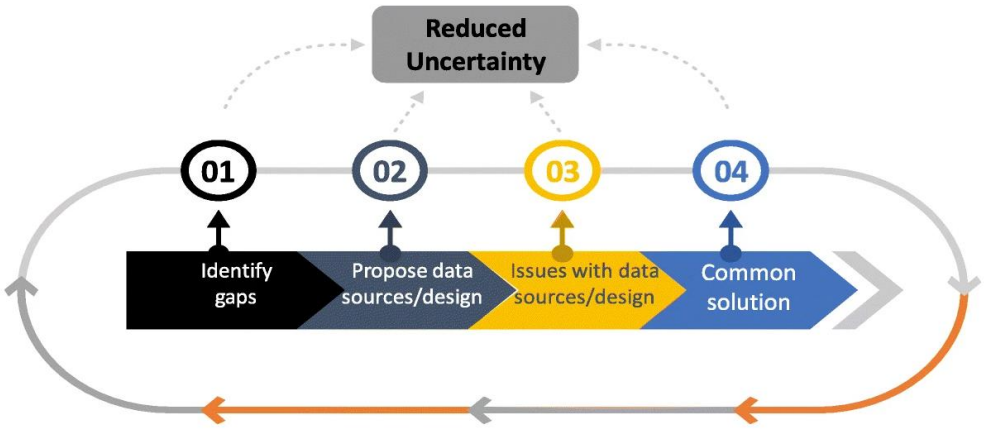


Potential evidence gaps need to be identified by the sponsor early, and agreement on timing and type of evidence needed to fill such gaps must be reached early enough to allow sufficient time to address the research questions.[44,45] It is especially important for the sponsor to deal with gaps in evidence for highly innovative, high cost drugs 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.[46] 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 (**Figure2**).[47] This framework provides guidance on how to determine the appropriateness and value of filling gaps in evidence with RWD throughout the lifecycle of a drug 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 TRUST4RD Tool

Source:[52]



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 create new questions to be answered.

The variety of evidence generated, as well as the amount of information derived from it, compel all stakeholders in drug development to recognize 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 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 industry, 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 originated by the evidence needed by different stakeholders.

Common stakeholder requirements/expectations are high quality data/information and reliability, access and understand the information.

Regulators request at 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 effectiveness and safety data from interventional and/or non-interventional studies in support of regulatory decisions. For example, 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. Appendix 1 provides a case study of the US FDA approval for fosdenopterin using externally controlled trials. (See case study A.)

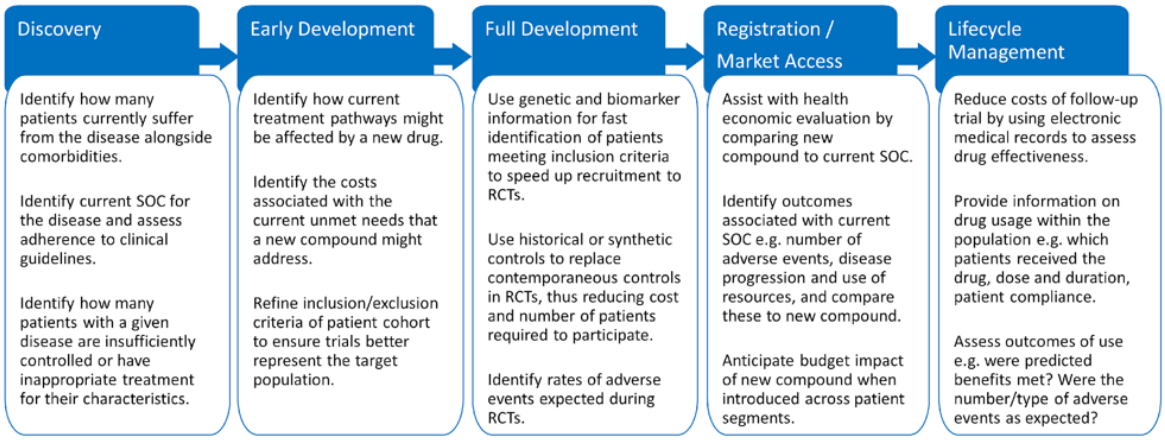
HTA requests cost effectiveness and budget impact analyses, in addition to the clinical efficacy and safety data. To meet HTA's requirements, sponsors provide cost estimates of the health state, QOL and utilisation of the health state, as well as economic models (e.g. SOC basis computed RCT results). 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 evidences that potentially satisfies the needs of both regulators and HTA bodies, the European Medicines Agency (EMA) offers consultations in parallel with the European Network for Health Technology Assessment (EUnetHTA), 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.[48]

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 HTA/payer provide evidence on diagnostic tools, optimal treatments and SOC, medical history and genetic information of the disease, potential drug-drug interactions.

1.4 RWE use in lifecycle of the development of medical products

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, can be used to complement them at various stages. Strategies that can facilitate and accelerate the drug development process are of high interest, and regulatory authorities have been evaluating the use of RWD across many stages of the drug development process.[49] **Figure 3** provides a summary of the various opportunities for RWE generation along the lifecycle.[50,51] Throughout this section, the typical applications of RWD in the product lifecycle will be further explored with the help of real-world examples.[52]

Figure 3: Examples of using real-world data (RWD) in the stages of the drug lifecycle
Source:[53]



1.4.1 RWE use in drug development phase

Examples of RWE use in a product's lifecycle can be found as early as during the compound selection of the target identification phase. The first step in the drug development process is the discovery of potential therapeutic agents, where researchers investigate the interactions among different molecules, genes, and proteins, with the goal to find novel targets, biomarkers, and compounds.[54] Some of these goals can be achieved using RWD applications. For example, in a recent review paper, 20 studies were identified that used RWD to facilitate drug discovery and clinical research. Among them, 16 identified or validated new phenotypes, disease markers, and biomarkers for patient identification and stratification.[55]

Within early research settings, RWD and RWE are being used to support the discovery of novel targets by identifying unmet medical needs, understanding disease epidemiology and characterising disease burden. They can focus 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 are being used to design and run clinical trials more efficiently by supporting:(1) better identification of target patient populations, (2) improved feasibility testing, (3) establishing the natural history of disease (particularly for rare diseases), (4) facilitating patient identification and recruitment, clinical site and country selection for global clinical trials, (5) identifying disease progression or mortality prognostic biomarkers to inform patient selection for trials (especially oncology drug development), (6) and accelerating 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 better 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 Plan Development.
- To better identify patients for participation in research programs, which speeds up the recruitment process and makes it more efficient. For example, well-managed databases based on EHRs allow queries leading to fast identification of patients meeting the recruitment criteria of an RCT.
- As input to make 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, enhancing representation of the population requiring access to the compound). For example, claims databases can show what the routine numbers of follow up visits and investigations are in daily practice and this practice can be mimicked in the pragmatic trial.

- 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 programs (from one protocol for one population with one drug to multiple combinations). RWD can be obtained through the cross-interrogation of multiple health care records containing genetic and biomarker information, which better enables the identification of target populations and therefore promotes inclusion diversity.
- To sometimes reduce the need for the recruitment of control patients to an RCT 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.
- RWD can be leveraged to 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 indispensable that the SOC is described as accurately as possible, and consider differences to be expected in different countries/regions.
- 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.
- To address 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 to.

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, further examples below:

- BAVENCIO® (avelumab) received accelerated approval by US FDA in 2017 for treatment of metastatic Merkel cell carcinoma and urothelial carcinoma, and conditional approval by EMA in 2017 for the treatment of Merkel cell carcinoma. These approvals were based on the assessment of a single-arm, open-label, Phase II study, JAVELIN Merkel 200.[56] In this study, historical controls based on McKesson`s iKnowMed electronic health care records and a German patient registry were used to characterise the natural history of Merkel cell carcinoma.[57]
- BLINCYTO® (blinatumomab) received accelerated approval by US FDA in 2014 and by EMA in 2015 for the treatment of relapsed/refractory Philadelphia chromosome-negative acute lymphoblastic leukaemia. These approvals were based on a single-arm, open-label, Phase II study. Data from this study were compared to a retrospective observational dataset obtained from national study groups and large treatment centres in Europe and the US.[58] A subsequent randomised Phase III trial (TOWER) run in 21 countries confirmed the efficacy of Blinatumomab in the relapsed/refractory setting as compared to SOC. Moreover, patients who received blinatumomab had better post-

985 treatment quality of life (QoL) compared to those on SOC[59]. Consequently, full approval
986 of the drug was granted.[60,61]

987 • In 2018, the US FDA granted BLINCYTO® (blinatumomab) a new indication for the
988 treatment of B-cell precursor acute lymphoblastic leukaemia (ALL) in first or second
989 complete remission with minimal residual disease greater than or equal to 0.1%. This
990 label extension was granted based on a propensity score analysis conducted to evaluate
991 the results of a blinatumomab multicentre, open-label, single-arm trial in comparison to
992 the historical data obtained from a retrospective OS that reviewed historical survival
993 data.[62,63]

994 • In 2019, Health Canada approved an expansion of the existing approved paediatric
995 indication for Prevnar 13 using RWD from the National Ambulatory Medical Care Survey
996 (NAMCS) and National Hospital Ambulatory Medical Care Survey (NHAMCS).[64] The
997 NAMCS provides information on the use of ambulatory medical care services in the US
998 based on visits to non-federally employed physicians and community health centres. The
999 NHAMCS provides information on the use and provision of ambulatory care services in a
1000 hospital emergency and by outpatient departments, with data compiled from visits to
1001 emergency departments, outpatient departments, and ambulatory surgery locations.[65]
1002 Based on the RWD provided, Health Canada approved the inclusion of acute otitis media
1003 in children six weeks to five years of age.[66]

1004 • In June 2020, Prolia® (denosumab) was approved by the National Medical Products
1005 Administration (NMPA) of the People's Republic of China as the first monoclonal
1006 antibody for the treatment of postmenopausal women with osteoporosis at high risk of
1007 fractures. The approval was granted with data from Prolia's global clinical trial program
1008 establishing favourable efficacy and safety, augmented by results from a RWD study
1009 confirming the effectiveness and safety of Prolia in clinical practice within Taiwan and
1010 Hong Kong.[67]

1011 For in-vitro diagnostic medical devices (IVDs), RWD can also play a crucial role in supporting
1012 regulatory decision making. Below we summarised examples of RWD use to support IVD
1013 regulatory intent.

1014 • One such submission is DEN170058, which relates to the MSK-IMPACT assay indicated
1015 as a next-generation sequencing-based tumour profiling test. It was supported by clinical
1016 data from an electronic medical record database of advanced cancer patients as part of
1017 routine workflow at Memorial Sloan Kettering Cancer Center. Retrospective analysis of
1018 these records provided evidence to support a pan-cancer claim, to validate a test cut-off,
1019 and to provide data on somatic mutation prevalence.

1020 • The marketing application for Placental Alpha Microglobulin-1 Immunoassay
1021 encompasses a total-product lifecycle example supported by clinical evidence in the form
1022 of patients' medical records. The sponsor submitted an observational clinical utility
1023 study of patients tested using the assay, for premarket clinical evidence and as a
1024 condition of approval.

1025 • A personal genome service from 23andMe supported a De Novo classification request
1026 using peer-reviewed, real-world literature as a primary source of clinical evidence for
1027 each of the ten conditions included in the Genetic Health Risk tests.

1028 • Information from the CFTR2 Database, a publicly maintained Next Generation
1029 Sequencing database, was used as the sole source of clinical evidence supporting a
1030 510(k) for both the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the
1031 Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay (Illumina, Inc. submission numbers
1032 K132750 and K124006).

1033 • Finally, an example of premarket paediatric RWE use is the SEEKER System (Baebies,
1034 Inc. submission number DEN150035) which was supported by a pivotal trial embedded
1035 in a state-run routine screening program, the Missouri State Public Health Laboratory,
1036 and Missouri Department of Health and Senior Services (MDHSS) Surveillance Program.

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

After market entry, RWD can help to:

- Provide evidence on the real-world usage of medicines, e.g. which patients received the drug at which dosages for what duration and patient adherence to treatments, especially for drugs known to have a high incidence of adverse events.
- Address safety-related or effectiveness-related questions (such as fulfilling a post marketing commitment), e.g. characterisation of an identified or potential risk, establish effectiveness of risk minimisation measure.
- Expand safety-related labelling (such as the warnings and precautions or dosing sections), e.g. identified need for further monitoring or visit to specialist to identify adverse effects early.
- Support the submission of marketing application renewals, if applicable, e.g. showing effectiveness in comparison to existing SOC, or global benefit-risk balanced in routine practice
- Support the conversion of a conditional approval to a full approval, e.g. additional safety and effectiveness established with RWD can support a confirmatory trial with limited sample size.
- Support a new indication or label extension, e.g. using pragmatic design features, extension of indication can be achieved
- Characterisation of special populations (older adults, etc.), e.g. access to a broader population without extensive inclusion/exclusion criteria to describe and assess safety and efficacy in specific sub-population.

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 programs 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, and to monitor off-label use of medications, which may be of value for both drug safety and drug repurposing (taking an existing drug 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 inform 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 may be beyond the scope of clinical trials.

1.5 Adapt good clinical practices concepts of data integrity to RWD

Regulatory agencies generally require the sponsor to submit data from RCTs in support of regulatory review. Yet the requirements for non-randomised studies are not entirely clear or consistent. Franklin et al. suggest that the submission of raw study data for regulatory submission of such studies seems imperative to address the concerns about the quality of both data and design in non-randomised RWE based on health care databases.[68, 69] The authors believe that sharing analytical programming code used for creating all analytic results, as well as code for cohort creation in the context of health care database studies, should be required of sponsors for regulatory submissions, and highly encouraged of all investigators in published literature.

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.[70] According to the draft guidance, the sponsor will retain and make available to the Agency upon request a log of any researcher or researchers who have significant involvement in the design or conduct of the study. Further, in the early stages of designing a non-interventional study intended for use in a marketing application, sponsors are requested to ensure that they are able to submit patient-level data for the RWD that have been analysed as part of the clinical study included in a marketing application.[71]

1.6 Evidence generation presentation and communication

In order 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 and reporting 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.

1.6.1 Existing guidance - Good Pharmacoepidemiology Practice (GPP)

Both the International Society of Pharmacoepidemiology (ISPE) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) have actively developed guidance for RWE studies.[72] Best practices include pre-specification of details of the study design and analysis plan and accountability for reproducible research.

1.6.2 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 guiding tool 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.[73] This information would increase health care decision makers' ability to effectively evaluate RWE studies. The recently published HARPER could be facilitated study protocol development and enhance transparency and reporting,[74]

In addition, to enhance transparency in RWD research, numerous public repositories exist for the registration of RWE protocols for future inspection, including the EU PAS Register[75], clinicaltrials.gov, and HSRProj. EU PAS register has also a source data repository available to also disclose information on the source of data.[76]

While transparency and disclosure are needed for evaluation, it is also the responsibility of the researchers to unambiguously communicate study results, 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 effect 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.[77] The inclusion of assessment of RWD in an effects table would make it explicit what "value" is added, and it would serve to build trust on reported RWE and establish the need for further investigations.

1.6.3 Cross-stakeholder interaction and collaboration

The need for discussion and consensus by multiple stakeholders around the acceptability of plans for generation of RWD/RWE has recently been highlighted.[78] For example, EUnetHTA was established to create an effective and sustainable network for HTA across Europe. EUnetHTA supports collaboration between European HTA organisations that brings value at the European, national, and regional level 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.[79] Since 2017, the EMA and European Network for Health Technology Assessment (EUnetHTA) have offered parallel advice services called Early Dialogues in order to provide a platform for such multi-stakeholder interactions.[80] This parallel consultation by regulators and HTAs to offer sponsors the opportunities for mutual understanding and problem solving between regulators and HTAs, the goal being to facilitate robust evidence generation for different stakeholders.[81] Another recent example of collaboration between regulators and HTAs is the formal recognition by the EMA of the EUnetHTA Registry Evaluation and Quality Standards Tool (REQuest).[82] 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 2018.83 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.

1.7 Engaging with regulators

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

1.7.1 US Food and Drug Administration

Mandated by the 21st Century Cures Act, the US FDA has developed a RWE Program and provided guidances on RWE use for regulatory decision making. RWE can be submitted to the US FDA in an Investigational New Drug (IND), Biologics License Application (BLA) or New Drug Application (NDA) submission or a meeting request, with a cover letter indicating that the submission contains RWE.[84] RWE submissions may come 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 submission 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.

1.7.2 European Union

The European medicines regulatory system is based on a network of around 50 regulatory authorities from the 31 European Economic Area countries (27 EU Member States plus Iceland, Liechtenstein and Norway), the European Commission and EMA. EU regulators use RWD analysis in post-approval on a regular basis, mostly to further characterise safety, but also that of

effectiveness.[85] During the pre-approval phase, the evidence generated from RWD has been seen to complement the evidence from RCTs.[86] There is however increasing interest in the use of RWD to support regulatory decision making across the product lifecycle.[87,88] Scientific advice[89] 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 program to provide parallel scientific advice (PSA) to sponsors. The EMA also offers consultations in parallel with the EUnetHTA 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 MA 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, high value seen in fast access to medicine or the infeasibility of performing a RCT or other challenges of following the traditional drug development pathway.[90]

1.7.3 General RWE landscapes in various countries

Australia

The Therapeutic Goods Administration (TGA) recently commissioned a review into their usage of real world evidence (and patient reported outcomes) in the regulation of medicines and medical devices.[91] The review found that there is ambiguity surrounding the usage of RWE and PROs, which potentially limits its adoption and that the stakeholders recommend that TGA improve their communication about how the TGA accept and use RWE and PROs.

The actions TGA have proposed as a response include creation of a central point for information about RWE and PROs on the TGA website, clarification of related definitions, requesting applicants to document why and where RWE and PROs have been included in the application and their purpose for inclusion as well as communicating when RWE and PROs are used in making regulatory decisions.

TGA is also to consult on relevant guidance for the use of RWE and PROs as evidence for the regulation of medicines and medical devices, covering generation of data (for inclusion in the dossier), and utilisation in evaluating the application. TGA will continue to learn from international sources for generation of RWE and PROs to maximise alignment with international regulator practices and aims to better understand how TGA might support the enhanced use of RWE and PROs into the future. This may include providing advice to potential applicants and designers of RWE and PROs programs intended for regulatory use, and the use of RWE and PROs for medicine regulation pathways such as orphan or provisional medicines, or for repurposing of medicines.

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 different stakeholders, such as academic institutions, the pharmaceutical industry, and regulators. At these meetings, the 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 drug development; data analysis driven by artificial intelligence in healthcare settings; opportunities and challenges of RWE studies; and perspectives of medical professionals and industry in relation to RWE.[92]

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 drugs aimed at treating rare diseases and serious and debilitating conditions. They are, 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; design of studies that include RWD; and others.[93]

This communication can be established through the following existing channels: pre-submission meetings for scientific advice (available for the drug 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 industry, but also by citizens and other government departments that are interested in seeking clarity from the Agency).[94]

Current strategies will contribute to the improvement of the current model of generating information, focusing on the subject/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 that is required for the continuity of subject/patient care. RNDS information may be valuable for epidemiological, statistical, research, and regulatory purposes.

In order to encourage the interoperability of health data through publication in a machine-processable format and promoting the continuous improvement of the quality of the data made available, Anvisa also developed an inventory of the databases under its custody to provide public knowledge about these databases maintained by the Agency. This initiative is called Anvisa's open data plan.[95] 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.[96]

Canada

The 2022-2023 Plan of Health Canada lists as its core responsibility to protect and promote health. Health Canada works with domestic and international partners to assess, manage and communicate the health and safety risks and benefits associated with health and consumer products, food, chemicals, pesticides, environmental factors, tobacco and vaping products, cannabis, and controlled substances. This focus includes to apply RWE in support of regulatory decisions to improve the post-market oversight of prescription drugs and medical devices in Canada, particularly those that treat rare diseases, as well as to inform decision making for COVID-19 drugs, vaccines and medical devices.[97] It will develop additional guidance on using RWE and will finalise a strategic plan with the CADTH and Quebec's INSSS to further align RWE use across the drug lifecycle. The goal is to improve the accessibility, affordability, flexibility and appropriate use of drugs in Canada.

Japan

The Pharmaceuticals and Medical Devices Agency (PMDA) uses RWD/RWE mainly for safety assessment in the post-approval setting.[98] The PMDA launched the Medical Information for Risk Assessment Initiative (MIHARI) project in 2009 with the aim of strengthening post-approval safety measures for pharmaceuticals.[99] In the MIHARI Project, PMDA has conducted safety assessments of drugs using pharmaco-epidemiological methods, with secondary use of electronic medical information that hospitals enter and accumulate for the purpose of routine

medical care, such as data contained in 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[100,101,102] and MID-NET[103,104,105] a reliable and valuable database operated and managed by the PMDA in Japan.[106] Some of those results have led to actual safety measures such as a revision of precautions of the package insert in Japan.[107,108] At the same time, to further improve post-approval pharmacovigilance in Japan, the GPSP (Good Post Marketing Study Practice) ordinance that set reliability standards for post-approval study conducted by the MAHs after drug approval were revised in 2017.[109] 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 construction of novel clinical development methodologies, more specifically, to construct the disease registry system and thereby proceed with construction of the clinical innovation network (CIN) that develops clinical development infrastructure based on the disease registry information.[110] Since then, joint industry-academia research-and-development projects that utilise the registries have been supported.

The registry utilisation for evaluating safety and efficacy of drugs and medical devices was clarified in the conditional accelerated approval system for drugs and medical devices, which has been started in 2017.[111,112] 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,[113] and point to consider for assurance of the reliability of utilisation of registry data as approval applications.[114]

In addition, the PMDA has started activities of Projects Across Multi-Offices, RWD Working group in April 2021, and discuss all subjects on RWD comprehensively including general principles on RWD utilisation and data reliability in regulatory settings.[115]

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 NMPA 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 drug review, indication expansion, post-approval evaluation, and R&D of traditional Chinese medicine. Following the publication of that guidance, a technical guideline on the development and review of drugs for children was released in September 2020 by the Centre for Drug Evaluation, an affiliated institution of the NMPA. Besides drugs, RWD are also used in the clinical evaluation of medical devices, for which a technical guideline was published by the NMPA in November 2020.[116]

References

[19] Definition from techopedia. What is interoperability? (Website accessed 28 October 2022).

[20] 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

[21] 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

[22] 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

[23] 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

[24] The United States Food and Drug Administration (FDA). Framework for FDA's Real-World Evidence Program. (PDF accessed 15 October 2022).

[25] 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).

[26] 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).

[27] 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).

[28] 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).

[29] 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).

[30] 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).

[31] 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).

[32] 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).

[33] European Medicines Agency (EMA). Scientific Guidance on Post-Authorisation Efficacy Studies. (Website accessed 15 October 2022).

[34] European Medicines Agency (EMA). Guideline on Registry-Based Studies. (Website accessed 15 October 2022).

[35] Heads of Medicines Agencies (HMA). HMA-EMA Joint Big Data Taskforce Phase II Report: 'Evolving Data-Driven Regulation'. (PDF accessed 15 October 2022).

[36] Heads of Medicines Agencies (HMA). Priority Recommendations of the HMA-EMA Joint Big Data Task Force. (PDF accessed 15 October 2022).

[37] 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)

[38] 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

[39] International Society for Pharmacoepidemiology (ISPE). Strategic Initiatives: RWE for Regulatory Decision Making. (Website accessed 15 October 2022).

[40] Canada's Drug and Health Technology Agency (CADTH). Use of Real-World Evidence in Single-Drug Assessments Environmental Scan. (Website accessed 15 October 2022).

[41] NICE real-world evidence framework. (Website accessed 23 May 2023)

[42] 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

[43] 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)

[44] 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

[45] 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

[46] 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 Ttechnology Assessment in Health Care. 2020;36(4):459-468. https://doi.org/10.1017/S026646232000063X.

[47] 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

[48] EMA Parallel joint scientific consultation with regulators and health technology assessment bodies. (Website accessed 2 June 2023)

[49] 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).

[50] Orpha Strategy. Patient-Centric Real-World Evidence. (Website accessed October 2022)

[51] Bate A, Juniper J, Lawton AM, Thwaites RM. Designing and incorporating a real world data approach to

[50] 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.

[52] 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).

[53] 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.

[54] The United States Food and Drug Administration (FDA). The Drug Development Process. (Website accessed October 2022).

[55] 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

[56] US National Library of Medicine (NIH). Avelumab in Participants with Merkel Cell Carcinoma (JAVELIN Merkel 200). (Website accessed October 2022).

[57] Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: the changing landscape of the regulatory approval process. Pharmacoepidemiology and Drug Safety. 2020;29(10):1201-1212. https://doi.org/10.1002/pds.4932

[58] Gökbuget N, Kelsh M, Chia V, Advani A, Bassan R, Dombret H, et al. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. Blood Cancer Journal. 2016;6(9):e473. https://doi.org/10.1038/bcj.2016.84

[59] Burt R, Warcel D, Fielding AK. Blinatumomab, a bispecific B-cell and T-cell engaging antibody, in the treatment of B-cell malignancies. Human Vaccines & Immunotherapeutics. 2019;15(3):594-602. https://doi:10.1080/21645515.2018.1540828

[60] Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: the changing landscape of the regulatory approval process. Pharmacoepidemiology and Drug Safety. 2020;29(10):1201-1212. https://doi.org/10.1002/pds.4932

[61] Amgen. FDA Grants Full Approval for BLINCYTO® (blinatumomab) to Treat Relapsed or Refractory B-cell Precursor Acute Lymphoblastic Leukemia in Adults and Children. (Website accessed 15 October 2022).

[62] The United States Food and Drug Administration (FDA). FDA Briefing Document: Oncologic Drugs Advisory Committee Meeting, BLA 12557 S-013 Blincyto (blinatumomab). (PDF accessed 15 October 2022).

[63] Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: the changing landscape of the regulatory approval process. Pharmacoepidemiology and Drug Safety. 2020;29(10):1201-1212. https://doi.org/10.1002/pds.4932

[64] Pfizer Canada ULC. Product Monograph: Prevnar 13. (PDF accessed 15 October 2022).

[65] Centers for Disease Control and Prevention (CDC). National Center for Health Statistics. Ambulatory Health Care Data. (Website accessed 15 October 2022).

[66] Zhou X, de Luise C, Gaffney M, Burt CW, Scott DA, Gatto N, et al. National impact of 13-valent pneumococcal conjugate vaccine on ambulatory care visits for otitis media in children under 5 years in the United States. International Journal of Pediatric Otorhinolaryngology. 2019;119:96-102. https://doi.org/10.1016/j.ijporl.2019.01.023

[67] Storm NE, Chang W, Lin TC, Lange JL, Bradbury B, Critchlow CW, et al. A novel case study of the use of real-world evidence to support the registration of an osteoporosis product in China. Therapeutic Innovation & Regulatory Science. 2022;56:137-144. https://doi.org/10.1007/s43441-021-00342-4

[68] 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

[69] 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

[70] 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).

[71] 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).

[72] International Society for Pharmacoepidemiology (ISPE). Resources: Guidelines for Good Pharmacoepidemiology Practices (GPP). (Website accessed 15 October 2022).

[73] 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

[74] 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

[75] 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).

[76] European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP). ENCePP Resources Database. (Website accessed 15 October 2022).

[77]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).

[78] 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

[79] European Network for Health Technology Assessment (EUnet HTA). About EUnetHTA. (Website accessed 15 April 2023).

[80] European Medicines Agency (EMA). Parallel Consultation with Regulators and Health Technology Assessment Bodies. (Website accessed 15 October 2022).

[81] European Network for Health Technology Assessment (EUnetHTA). Guidance on parallel EMA/EUnetHTA 21 Joint Scientific Consultation. (PDF accessed 15 October 2022).

[82] European Medicines Agency (EMA). Guideline on Registry-Based Studies. (PDF accessed 15 October 2022).

[83] Government of Canada. Notice to Industry: Aligned Reviews between Health Canada and Health Technology Assessment Organizations. (Website accessed 15 October 2022).

[84] 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).

[85] European Medicines Agency (EMA). Scientific Guidance on Post-Authorisation Efficacy. (PDF accessed 15 October 2022).

[86] European Medicines Agency (EMA). Regulatory Perspective on Real-World Evidence (RWE) in Scientific Advice. (PDF accessed 15 October 2022).

[87] 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

[88] 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

[89] European Medicines Agency (EMA). Human Regulatory: Scientific Advice and Protocol Assistance. (Website accessed 15 October 2022).

[90] 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

[91] Australian Government. Department of Health and Aged Care. Review of real world evidence and patient reported outcomes. (Website accessed 15 October 2022).

[92] 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).

[93] 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).

[94] 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.

[95] *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*

[96] *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*

[97] *Government of Canada. 2022-23 Departmental Plan: Health Canada. (Website accessed October 2022).*

[98] *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*

[99] *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*

[100] *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*

[101] *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*

[102] *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*

[103] *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*

[104] *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*

[105] *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*

[106] *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*

[107] *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*

[108] *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*

[109] *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).*

[110] *Japan Revitalization Strategy 2016. (PDF accessed on 2 May 2023).*

[111] *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).*

[112] *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).*

[113] *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).*

[114] *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).*

[115] *Pharmaceuticals and Medical Devices Agency (PMDA). Projects Across Multi-Offices in PMDA. RWD WG. (Website accessed 2 May 2023).*

[116] *Zhong J, Zhang J, Fang H, Liu L, Xie J, Wu E. Advancing the development of real-world data for healthcare research in China: challenges and opportunities. BMJ Open. 2022;1;12(7):e063139. http://dx.doi.org/10.1136/bmjopen-2022-063139*

1312

Chapter 2: Real-world data sources

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 in terms of capturing information arising from the provision of health care, are useful to evaluate the safety and effectiveness of the 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 and, for the purpose of this document, will be called traditional data sources.

The introduction of new technologies such as those related to virtual care and the increased use of mobile devices has provided new sources of different types of information that can be generated with unprecedented volume, speed and complexity and require a different set of data management and analytical methods. Although the 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 Health care databases

RWD from health care databases, including insurance claims, EHR, and registry databases have been used for decades, mainly for safety evaluation, risk management and to support benefit-risk evaluation of medicinal products.[117,118] 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 occurring during the follow-up allow for the estimations of “true” risks. “True” in the sense that such estimates will be more accurate than those from SRSs, for example, where the size of the exposed population at risk is not available.
- 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 a longer follow-up time. Thus, they are more suitable to identify 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.

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 to answer 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. The linkage can be direct or in the form of a comparison, depending on the settings of databases being attempted to link. There are several issues to consider when directly linking data including confidentiality and ethical issues. Data fusion is a method that can use multiple data sources without direct data linkage. The estimated values of other data sources can also be used. An example of data fusion is using data from the cancer registry to determine the incidence per population, and then using data from the transplant registry to investigate whether the mortality rate of survivors is higher compared to the general population if they are long-term survivors after, for example, cancer or transplantation. Different data sources are used, but not exactly via data linkage.

Registries usually have information only for patients exposed to specific drugs and/or experiencing certain diseases. Another limitation is 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. Availability of information for certain sub-groups of important patient populations (for example, older adults, children, and pregnant women) may also vary. Finally, 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.[119,120]

Given these limitations, when performing new studies, it is important to involve or consult with the parties who are close to the development of those health care databases, or who have had experience in using them, 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 the potential for bias due to 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. Therefore, if a p-value (the probability of obtaining test results at least as extreme as the result actually observed, under the assumption that the null hypothesis is correct) is used to measure the statistical significance of an association, should it be adjusted to address multiple analyses? Some suggest adjustment is not necessary because it is not a clinical trial[121] while others prefer some kind of adjustment.[122]

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)[123,124] but less commonly for signal detection with no a priori hypothesis.[125] Besides the challenges already mentioned above, the use of the same database both for signal detection and signal evaluation presents another challenge.[126] Some suggest that signal detection (or hypothesis generation) should be done independently from signal evaluation (or hypothesis testing) in a different data source.[127] Others suggest that the two could be done in the same databases as long as the methods of analyses are different.[128,129]

Although the use of health care databases for RWD studies on benefits (effectiveness) has been limited and more controversial[130] there has been a lot of discussion on how they can be used as part of regulatory decision making. Many of the 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.[131]

2.1.2 Ad-hoc data collection

When the existing data sources are not suitable to answer the questions at hand, either due to the lack of information or differences in the study populations, a new RWE with ad-hoc data collection is needed. For example, most existing data sources lack information connecting mothers and their babies and, therefore, are not suitable to evaluate the associations between exposure to medicinal products used during pregnancy and the pregnancy outcomes. The creation of a pregnancy registry can be done on an ad-hoc basis and designed specifically to test a particular set of hypotheses.

Another example is when the US FDA issued guidance to address cardiovascular safety issues that might be related to new anti-diabetes drugs among patients with type 2 diabetes.[132] This guidance led to new randomised or non-randomised studies being performed that extended the evidence from pre-approval clinical trials to post-marketing real-world settings. These studies can be used to address safety issues or to evaluate the real-world effectiveness of not only anti-diabetes drugs but also other products.

Real-world studies with ad-hoc data collections can be longitudinal, such as in examples mentioned above, or cross-sectional. A drug utilisation study is usually done cross-sectionally to measure the effectiveness of risk minimisation actions to limit the use of a drug, for example among contra-indicated patients. For this purpose, a drug utilisation study is usually done repeatedly in the same population, before and after the minimisation measure is implemented.

While RWD with ad-hoc data collections share the same strengths as health care databases, they also have an additional advantage. Ad-hoc data collection is performed specifically to answer a set of questions and, therefore, potentially more effective in answering those questions. Despite the strengths, RWD sources with ad-hoc data collection also have limitations. The study subjects (for example, patients or HCPs) participate in the study on a voluntary basis, and it may take a long time to accrue enough subjects in the study. 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 drugs, 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.

2.1.3 Federated systems

The availability of many different RWD data sources presents a unique opportunity to perform the same study using different sources. Consistency of the results, or lack thereof, will help to understand the research question being evaluated better by potentially enlarging the sample size, including diverse patient population, enriching health care data, and prolong study follow-up time. There are two approaches to performing a study using multiple data sources: by pooling the raw data together or by analysing the data separately and then combining the results using, for example, a meta-analysis. The former may be problematic, as data sources are originally built and developed with different purposes and formats, making the pooling of the raw data very difficult, if not impossible. The latter, called sentinel system, is more appropriate and there are currently a few systems available. See the CIOMS Working Group X report on "Evidence Synthesis and Meta-Analysis for Drug Safety".[133]

Different RWD sources consist of data collected for different purposes and with different designs, using different formats, and utilising different codes for diseases, conditions and medicinal products as well as devices. In a sentinel system, these different codes are harmonised and standardised into a single system. Such a standardised system is called a Common Data Model (CDM). The CDM was first developed by The Observational Medical Outcomes Partnership (OMOP), a public-private partnership established "…to inform the appropriate use

of observational health care databases for studying the effects of medical products"134 This partnership has ended, and the legacy has continued with the Observational Health Data Sciences and Informatics (OHDSI), one of the sentinel systems available to date. OHDSI currently uses OMOP CDM version 5.4.[135] The key goal of OHDSI is to facilitate large-scale observational research studies by leveraging diverse sources of real-world health data, such as electronic health records, claims databases, and registries. By transforming and mapping these heterogeneous data sources into the OMOP CDM, OHDSI enables researchers to conduct studies on a massive scale and combine data from multiple institutions and countries.

The European Health Data & Evidence Network (EHDEN) (reference here) is another initiative in Europe related to OHDSI. EHDEN is a public-private partnership that aims to accelerate the generation of real-world evidence (RWE) and it was funded by the Innovative Medicines Initiative and the EU. While EHDEN focuses primarily on Europe, OHDSI is an international collaborative community with a global reach, both using the OMOP CDM standards and analytical tools.

Another system using a CDM approach is the US FDA Sentinel System. The US FDA Sentinel System is an active safety surveillance system for US FDA-regulated medical products, using a distributed database of primarily electronic claims data collected as part of 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 that are stored in a common data model at each site.[136]

The US FDA Sentinel CDM specifies how data are stored, structured, and labelled for all data partner sites. Many organisations contribute to the Sentinel Distributed Database and adhere to a CDM to assemble patient - level files from their source data. Each participating organisation designed a process to extract, transform, and load its source data, applying the common data model to create the Sentinel Distributed Database. Organisations adhere to clinical coding standards, such as ICD-9 and NDC codes; locally developed codes are occasionally used, and the CDM accounts for that coding variability. CDM allows various latency and frequency with which data partners can refresh the data.

The current version of US FDA Sentinel CDM (8.0.0) included 16 tables representing specific data domains that are available in administrative and claims data, such as demographics, dispensing and encounter data.[137] The table structure meets the need for data access while preserving the granularity and nature of the source data.[138] Unique person identifiers allow linkage across the tables to provide a comprehensive, longitudinal view of patient care. The CDM can be expanded to accommodate new data domains, typically through the addition of new tables to the existing model.

In the US FDA Sentinel system, to ensure conformance to CDM specifications, the completeness and content of each variable in each table are examined at regular intervals, as well as the logical relationship and integrity of data values within and across variables and within and across tables. Finally, the consistency of data distributions is examined over time and across data partners.

The advantage of using a CDM lies in the fact that investigators can pull multiple data sources together into one unified data set (either centralised or distributed) that could provide larger sample size, broader patient populations, and enriched details in healthcare utilisations. However, the use of CDM might result in 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 (e.g. US FDA Sentinel, the National Patient-Centered Clinical Research Network (PCORnet®), and OHDSI), 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.[139]

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP®), a network coordinated by the EMA, is another sentinel system.[140] Different from OHDSI and the US FDA Sentinel, ENCePP is a network of investigators using different European RWD sources separately without a CDM. Besides ENCePP, there is another system in Europe that utilises health care databases across the EU, the Data Analysis and Real World Interrogation Network (DARWIN EU®). Established by EMA, DARWIN EU is "a coordination centre to provide timely and reliable evidence on the use, safety and effectiveness of medicines for human use."[141] Unlike in the case of ENCePP, in the DARWIN EU project (or initiative or coordination centre, but not system), the databases are analysed separately in a federated network model using the OMOP CDM standards and analytic tools.

2.1.4 Other sources

Another important RWD source is SRSs such as the US FDA Adverse Event Reporting System, US FDA Vaccine Adverse Event Reporting System, WHO Vigibase, and EudraVigilance, a system for managing and analysing information on suspected adverse reactions to medicines which have been authorised or 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[142,143] and will remain so for the foreseeable future.

A SRS consists of individual case study reports spontaneously reported by patients, HCPs and other reporters (such those who become aware of the cases and then report them to the producers or market authorisation holders of the products). 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 no follow-up on individual patients, which is critical in the evaluations of associations between medicinal products and events. Other well known weaknesses include underreporting (not all events are reported), stimulated reporting (the reporting of events can be increased by factors like publicity), differential reporting (events related to certain drugs may be more likely to be reported than events related to others), and poor data quality in terms of validity and quantity (e.g. the same event resulting in multiple reports).[144,145,146] Another limitation is the Weber effect, in which there is a gradual increase in reporting within early years after launch.[147] A more recent study suggests that the Weber effect does occur within newer, more modern adverse events reporting systems.[148]

Despite their limitations, SRSs play a key role in identifying and addressing safety issues. One of the SRS' strengths is the large amount of data that allows for 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 to be useful in addressing the issues appropriately. Another strength is that it is 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 or benefit 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 to estimate 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 medicinal 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 becoming connected in order to obtain data for real-time patient assessment or for reporting purposes in clinical trials.

At the same time, other important forces are converging, such as improved access to genomics data, the adoption of machine learning models for data analysis, and the move toward 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 to fulfil the promise of improved safety signal detection, personalised medicine, impactful clinical research, reduced health care costs and population health management.

2.2.1 Biosensor data

This is a RWD source 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 sensor to detect atrial fibrillation and other arrhythmias.[149] This will allow for a more effective monitoring especially in a challenging time such as the lockdown during COVID-19 pandemic, for which the US FDA issued a 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."[150]

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 "good 'insert activity' practices (GxP)" validated devices, which permit processing of data that need to be 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 crystallise its use.

2.2.2 New sources of data for collection of patient reported outcomes

Data that are reported by patients may include information about patient preferences, patient experiences and patient health outcomes for example self-reported joint pain/mobility scores, changes in the severity of a dermatologic condition. Patient preferences enable the consideration of the relative desirability or acceptability of different alternatives or choices

among interventions or endpoints, or alternative care pathways. Patient experience data considers perspectives, needs and priorities related to a disease or condition. Patient reported outcomes (PROs) refer to data concerning the patient's health condition or status, from the patient's view rather than the view of a HCP or test. 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 EHR plus ancillary data using a specific methodology and a common data model so that outcome and exposure definition are robust

The volume and variety of health care-related data which has been added to the electronic medical records has continued to grow in recent years and has been considerably enhanced by the recent availability of 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 not viable without implementation of a curation process. A comparison[151] between several EHRs with and without curated data showed 67% concordance when relying on structured data alone versus 97.5% concordance among curated records. Another challenge is that in some cases the 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.[152,153] We need to be clear about definitions such as overall survival, disease-free survival, objective response rate, complete response rate, progression-free survival. Even with curated data, different approaches can lead to different results. For example, different methods of calculation of progression free survival in breast cancer have been shown to shift the median time to progression by months. Such issues have been addressed using an oncology specific common data model such as mCODE[154] and its implementation in HL7 FHIR.[155] 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 Data in the form of text or images coming from radiology information systems

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 the use of non-digitised 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 an ability for personalised image interpretation (e.g. by a radiologist caring for the patient), discovery of new imaging markers, and wider utilisation of data 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.

The identification of an adverse event such as pulmonary embolism (PE) can be readily done using computed tomography angiography (CTA), which is the test of choice. Nonetheless the actual rate of positivity is rather low (10%) due to the difficulty of selecting patients with a high pre-test probability. However, machine learning models using RWD from large numbers of patients concerning clinical, lab and other radiological information (e.g. chest x-ray) could presumably be used to risk-stratify new patients and increase the CTA positivity rate.

2.2.5 Data in the form of PDF text/images, structured lab output data, coming from the Laboratory Information System

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. Lab data can aid decision making and help to measure endpoints, outcomes, or exposures either alone or when included in algorithms, thus helping the data curation process.

Although lab data has been routinely used as secondary data for research purposes, the laboratory information system has been 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 lab data requires investigators to invest considerable time cleaning the lab dataset. While there are efforts in progress to improve the lab information system such as by using or mapping Logical Observation Identifiers Names and Codes (LOINC)[156] codes (see Chapter 1 on Uses of real-world evidence for decision making during the product lifecycle), coverage is not perfect ranging from 73% to 90% for a reference laboratory which handles both common and specialised tests.[157]

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

2.2.6 Data from any type of structured genomics investigation (full genome scan)

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.[158] The broad distribution of open datasets has required the creation of large-scale dataset repositories such as the National Center for Biotechnology Information (NCBI), Sequence Read Archive (SRA), European Nucleotide Archive (ENA), GenBank, and Protein Data Base (PDB). 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[159] sites.

The difficulty for an investigator is the need to combine genomics data with phenotype data. There are few cohorts / registries with such merged data being available for analysis. One example is the Genetic Epidemiology Research on Aging (GERA), which involves 78,000 subjects and 55 billion bits of genetic data, that is linked with comprehensive longitudinal electronic medical records as well as survey data on participant's health habits and background. Merged phenotype/genotype databases provide a unique opportunity to perform advanced analytics concerning safety not only in clinical trials, but also for post-marketing studies. However, evaluation of drug 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 issues (see Chapter 4 on Ethical and legal issues in using RWD). 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. Furthermore, genomic privacy is particularly problematic since the genome carries more individual data than one's credit card transactions. The Global Alliance for Genomics and Health (GA4GH) has worked to develop ways to balance the concerns of individual privacy and the social benefits of data sharing.[160]

2.2.7 Data from social media

The interest in the potential use of data from social media for safety surveillance has been increasing in the last decade. One study showed that there was a concordance of the numbers of adverse events from twenty SOCs mentioned with medical products in Twitter and those reported in the US FDA Adverse Event Reporting System.[161] This concordance does not

necessarily mean that social media data is a reliable source 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-Recognizing 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."[162] Another study showed that, if the data were limited to patient groups, these signal detection methods performed better with the sensitivity ranging from 29 to 50.6% and the specificity from 86.1 to 95.5%.[163] 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 to prioritise which events should be evaluated further using more reliable data sources. In addition, it could also be used to evaluate the reasons as to why people reject vaccines or stop using medications.

Finally, as there has been a growing recognition of the importance of incorporating patients into decision making throughout the lifecycles of drugs and medical devices, social media is an important RWD source to obtain patient needs and perspectives, including patient preferences and patient reported outcomes. [164,165,166,167,168,169]

References

[117] *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

[118] *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

[119] *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

[120] *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

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

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

[123] *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

[124] *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*

[125] *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

[126] *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

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

[128] *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

[129] *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

[130] *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*

[131] *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*

[132] *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).*

[133] *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*

[134] *Observational Health Data Sciences and Informatics (OHDSI). The Observational Medical Outcomes Partnership (OMOP). (Website accessed 15 September 2022).*

[135] *Observational Health Data Sciences and Informatics (OHDSI). OMOP Common Data Model. (Website accessed 2 June 2023).*

[136] *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*

[137] *The United States Food and Drug Administration (FDA). Sentinel Initiative. (Website accessed 15 September 2022).*

[138] *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*

[139] *The Office of the National Coordinator for Health Information Technology. Common Data Model Harmonization. (PDF accessed 15 September 2022)*

[140] *European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). What is ENCePP? (Website accessed 7 April 2023).*

[141] *European Medicines Agency (EMA). Data analysis and Real World Interrogation Network (Darwin EU). (Website accessed 15 September 2022).*

[142] *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. doi.org/10.2165/00002018-200528110-0000*

[143] *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*

[144] *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*

[145] *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*

[146] *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*

[147] *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.*

[148] *Hoffman, Keith B et al. The Weber effect and the United States Food and Drug Administration's Adverse Event Reporting System (FAERS): analysis of sixty-two drugs approved from 2006 to 2010. Drug Safety. 2014;37(4):283-294. https://doi:10.1007/s40264-014-0150-2*

[149] *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).*

[150] *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).*

[151] *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*

[152] *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*

[153] *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 Health. 2019;22(6):739-749. https://doi:10.1016/j.jval.2019.01.014*

[154] *minimal Common Oncology Data Elements (mCODE). (Website accessed 15 September 2022).*

[155] *minimal Common Oncology Data Elements (mCODE). Implementation Guide. (Website accessed 15 September 2022).*

[156] *Health Information Technology Standards Committee. Letter to: Dr. Farzad Mostashari, National Coordinator for Health Information. 2011. (PDF accessed 15 September 2022)*

[157] *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*

[158] *O'Brien SJ. A decade of GigaScience: A perspective on conservation genetics. GigaScience. 2022;11,1-3. https://doi.org/10.1093/gigascience/giac055*

[159] *Github.com (Website accessed 29 September 2022).*

[160] *Knoppers BM. International ethics harmonization and the global alliance for genomics and health. Genome Medicine. 2014;6(2):1-3. https://doi.org/10.1186/gm530*

[161] *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. doi.org/10.1007/s40264-014-0155-x.*

[162] *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*

[163] *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*

[164] *European Medicines Agencies (EMA). The patients' voice in the evaluation of medicines. (PDF accessed 15 September 2022).*

[165] *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*

[166] *The United States Food and Drug Administration (FDA). Center for Biologics Evaluation and Research (CBER). (Website accessed 15 September 2022).*

[167] *U.S. Department of Health and Human Services (HHS). (Website accessed 15 September 2022).*

[168] *The United States Food and Drug Administration (FDA). Center for Devices and Radiological Health (CDRH). (Website accessed 15 September 2022).*

[169] *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*

1723

1724

1725

Chapter 3: Key scientific considerations in regulatory real-world evidence generation

The different sources of RWD (see Chapter 2 on Real-world data sources) derive from different settings, including both primary and secondary data sources, and RWE can be derived from different types of data for different types of research questions and hypotheses.

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 or any other medicine/vaccine related problem) is an important research field where RWE is used to address safety issues. Evidence on safety of drug therapy is incomplete at drugs' approval because clinical trials conducted are limited to their analysed patients' characteristics, sample size and study duration. Therefore, post-marketing surveillance is necessary to examine immediate and long-term effects, both safety and effectiveness, under real treatment conditions (e.g. considering incomplete medication adherence and the presence of comorbidities that may have been exclusion criteria of RCTs). However, 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. 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.[170]

3.1 Data source and data quality, integrity, transparency for data transformations, fitness for purpose

In the setting of traditional clinical trials or prospective/retrospective studies which collect data according to the research plan, data collection phase is included in the research, thus, data items to be collected and their definitions are designed prior to data collection.

In contrast, in RWE generation, existing RWD or existing database/platform is often used. RWD data sources described in Chapter 2 are often created for no-research purposes. This means that they may or may not be fit for a specific research purpose.

Therefore, it is important to have a good understanding of the characteristics of the data source and to make the necessary evaluations before using RWD for the purpose of RWE generation. 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- consideration points are discussed focused on data source types.

3.1.1 Scientific considerations for evaluation/selection of the database: evaluation of fitness for purpose

Scientific fitness for purpose is critical for selection of the database. The attributes of an RWD source have to be suitable for and relevant to the purpose of the study, including the size and representativeness of the study population and the availability of key variables. 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 subjects for the planned study is essential. If the population covered is a subset of data source, the variables to ascertain such population should be available, and the number of patients should be sufficient for the study objective. A particularly important point of consideration is generalisability, or external validity of the study population to the population to whom the evidence will be applied.

Especially in the secondary use of existing data, it is critical whether the key variables (exposure, outcomes/endpoints, demographic characteristics, and potential confounders) required to answer the clinical 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 for the possibility of additional 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 of time, attention should be paid to changes in the definition of common variables in the relevant area, including disease classification.

A single database may not be sufficient for a given research question and multiple databases may need to be used. If that is the case, the same principles of scientific considerations of fitness-for-purpose apply to each database. When multiple data sources are used, proper work processes in the data ingestion and harmonisation of datasets into a common data model are extremely important. Potential biases such as data availability bias and selection bias also need to be considered.

3.1.2 Feasibility considerations for evaluation/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 the 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 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 be started in the early phase in study planning.

Access to data

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

3.1.3 Quality considerations for evaluation/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 true copy, and accurate. The system and procedure to maintain data integrity (for example, how accuracy and consistency is assessed and who is responsible) is very important. At secondary use of existing data, evaluate the data lifecycle and its process in the perspective of quality management. 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.

Transparency for data transformations including mixed data sources

Investigators should check if the data transformation/data manipulation process is described in the lifecycle of the database or mixed data sources. Definitions to describe data

transformation/data manipulations need to be reviewed carefully as well as the data transformation/data manipulations process. When using mixed data sources, it is likely that a data transformation/data manipulation process is required to perform analyses. In this case, data transformation/data manipulation dictionary and its process need to be defined and be performed accordingly. This also ensures data traceability.

Regulations and good practice

It is necessary to follow the regulations set by the country or region for each purpose of utilisation. Quality assurance process by using site audit may be requested according to the regulations, but challenges including resources and access remain.

3.2 Study design and methods

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

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.

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, 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 disease (cases) to that of individuals without the disease (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.

Self-control case series studies

Self-control case series studies compare the occurrence of an event in an individual during a period when they are 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 haemoglobin A1c in type 2 diabetes), an intervention (e.g. surgical procedure), or a prescription for a new drug. 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.[171] 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.

Race and ethnicity

Constructs such as race and ethnicity merit additional care in the design and analysis of studies that will 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 23 years but the quality of reporting has not.[172] Many healthcare databases contain limited if any data on race/ethnicity and lack critical details regarding the way in which those data were collected. The measurement of race/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, one might think of race/ethnicity as a confounder and/or and effect modifier. The use of race/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[173]. Estimating differences in the effect of a treatment in specific populations based on race and/or ethnicity should only be done informed by an understanding of the local social, economic and institutional patterns that may influence health and healthcare. 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 that are 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,[174] researchers are advised to seek up-to-date expert guidance on measurement, analysis and reporting of race/ethnicity.

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 drug effectiveness and safety involves careful consideration of disease or condition factors and data sources.[175]

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, and disease-specific 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 case 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. Composite endpoints, which are composed of a series of items, 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. RWD collected for 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. Subjective assessment for clinical outcomes may also be considered for qualification for use in RWD studies.

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."[176] 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 are defined as "the capacity to perform the usual daily activities for a person's age and major social role", and often includes physical functioning, psychological well-being, and social role functioning.[177] 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)[178], SF-36 (measurement of 8 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),[179] 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).[180] Disease- or population-specific questionnaires, which may be more sensitive to symptoms that are experienced to particular subjects, 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 are outcomes that are used as substitutes for clinical outcomes, also referred to as intermediate endpoints, that are thought to predict clinical outcomes. Examples of surrogate outcomes include biomarkers, imaging findings, or laboratory tests that are thought to be associated with a particular disease or condition. Intermediate endpoints may be used to reduce the follow-up period required to obtain results, thus is more commonly used in clinical trials than in OSs. However, if the surrogate outcome does not reliably predict the occurrence of the clinical endpoint of interest, unhelpful or even harmful interventions can look beneficial.

d. Economic outcomes

These are outcomes that measure the economic impact of an intervention or exposure. Examples of economic outcomes include direct costs such as healthcare costs, indirect costs such as productivity loss, and quality-adjusted life years (QALYs).

Exposure definitions

Selecting the appropriate exposure definition is critical in the real-world assessment of drug effectiveness and safety. The section below details the three most common strategies (a-c), along with their strengths and weaknesses.

a. On-treatment exposure definition

The on-treatment exposure definition, also known as the as-treated exposure definition, follows patients from the start until the end of their treatment. Thus, events occurring during the follow-up period occur while patients are on treatment. This exposure definition inherently assumes that the drug has a reversible effect on the outcome (i.e. the effects disappear after treatment discontinuation). This exposure definition is well adapted for acute outcomes that are thought to be prevented or caused while exposed to a given drug (e.g. myocardial infarction, 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 severely affect its sensitivity.[181] 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 motivated by the frequency of the prescribing patterns (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 satisfied, 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.[182,183,184]

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

The as-started exposure definition, which is analogous to the intention-to-treat principle used in RCTs, follows patients from the start of their treatment until the end of follow-up, regardless of treatment discontinuation.[185] Compared with the on-treatment exposure definition, the as-started exposure definition is simple to implement, as gaps between prescriptions and grace periods are not considered. While this exposure definition is used in trials to maintain the balance achieved by the randomisation process, its use in real-world studies does not maintain or guarantee balance in patient characteristics. This exposure definition answers the clinical question of whether to initiate a drug 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 onsets, such as cancer incidence (especially if it is thought to have an irreversible effect on the drug). 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.[186]

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,[187] while having the advantage of dealing with complex exposure patterns. 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. Moreover, this definition is subject to time-dependent confounding if covariates are measured at baseline. This potential time-dependent confounding can be addressed using analytical approaches including marginal structural models.[188,189]

One commonly employed 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 nonrandomised RWD studies as emulations of a target RCT.

New-user vs. prevalent user definition

One way of emulating a trial is to conduct a new-user, active comparator design.[190] This process typically involves identifying an exposure of interest and an active comparator. Both the exposure and active comparator should be new users, which avoids prevalent user bias and ensures that patient characteristics are measured before the initiation of the exposures, avoiding adjustment for factors affected by the choice of treatment.[191] 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 drugs of interest at some point in the past, but not during the washout period (e.g. one year before cohort entry). While there is no clear consensus on the optimal length of the washout period, it should accommodate 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 nonrandomised 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 should consist of a drug 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.[192] Second, the use of an active comparator facilitates the 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, there are clinical situations where the question is whether we should initiate Exposure B versus continuing with treatment strategy A. This is an important question that is often addressed in trials. The comparator group consists of no active treatment or SOC (such as in the cardiovascular outcome trials of novel antidiabetic drugs). 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 who failed on Exposure A. The prevalent new-user design was specifically designed to address these real-world situations.[193]

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 lies in 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.[194] 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.[195] 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 recently implemented to assess the cardiovascular safety of aromatase inhibitors in women with oestrogen-positive breast cancer.[196] This study compared patients switching from tamoxifen to aromatase inhibitors with patients continuing treatment with tamoxifen.[197] 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 was investigated in several trials, and thus the prevalent new-user design emulated these trials.[198] This is distinct from another study using a new-user, active comparator design comparing new users of aromatase inhibitors with new users of tamoxifen;[199] that study assessed whether the upfront initiation of these drugs 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 make sure that the treatment and comparison groups are comparable. While in clinical trials this comparability can be achieved via randomisation, in RWD studies it can be achieved, among other approaches, by addressing the issue of confounders.

Confounding is one of the biggest challenges in working with RWE and plays an even more significant role when making statements about treatment effectiveness compared with safety. Confounding is present when the association between exposure and the outcome is disturbed by the presence of a third variable (the confounder). A variable is a confounder if it is associated with (1) the exposure, (2) the outcome, and (3) is not in the causal pathway between the exposure and outcome, or is not an intermediate variable.

- Confounding by indication

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

- Time dependent confounding

  Time-dependent confounding refers to confounders that change over time. In the case that information of a confounder in different points of time is available (such as body weight and laboratory values), this type of confounder can be addressed using analytical approaches including marginal structural models.[200,201]

Descriptions on bias and unmeasured confounding are provided in more detail in section 3.2.4 on Bias and unmeasured confounding. Statistical methods to improve comparability (e.g.

matching and adjusted analysis) are discussed in section 3.3 on Considerations for statistical analysis in RWD setting.

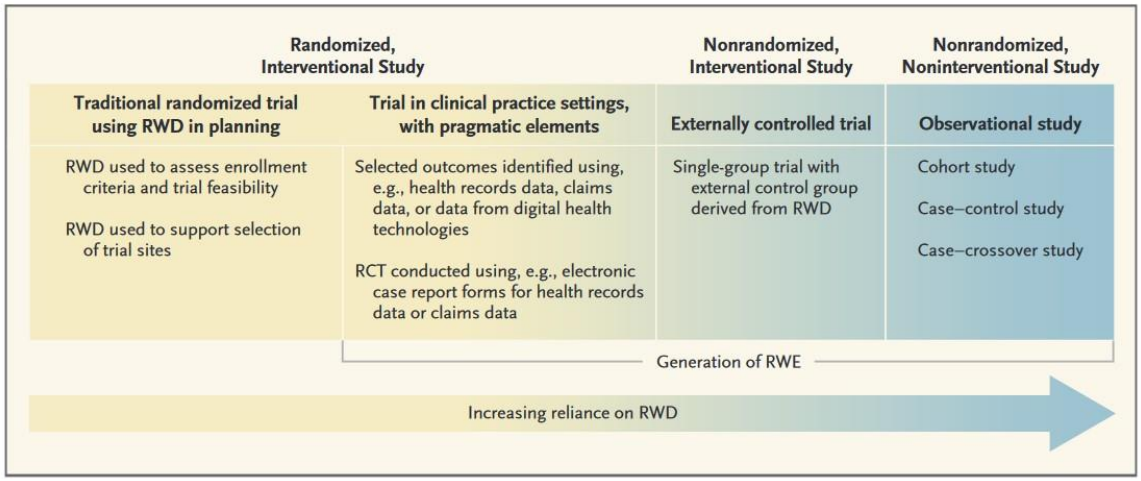
3.2.3 Study design considerations in context of RCTs

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, valid with high internal validity to support a causal inference. The reliability of RCTs is further supported by features that ensure an accurate assessment of trial eligibility, treatment exposure (intervention) and outcomes. These features include a well-defined, specific trial entry/exclusion criterion, 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 subjects are highly selected (may exclude older patients with comorbidities or concomitant medications) and sample sizes are small.

The following figure illustrates the various interventional and non-interventional study designs where RWD/RWE can be integrated into clinical evidence development of the effectiveness and safety of medical products during the entire product lifecycle.

Figure 4: Reliance on RWD in representative types of study design

Source:[202]



Traditional RCTs using RWD elements

Traditional RCTs are usually defined as an interventional research design in which one or more human subjects 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 CRFs that are separate from routine medical records. This helps to ensure 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 relevant data for the assessment of relevant outcomes (radiographic or results of exercise stress tests). For example, traditional RCTs of direct acting oral anticoagulants vs warfarin, double blinding limited the close monitoring of warfarin treatment to ensure it remained within the therapeutic range 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 (pRCTs)

These trials are largely thought to provide answers to important and relevant questions about the real-world effects of treatments in the post-approval routine clinical practice settings.[203] The degree of pragmatism varies and such studies typically run on a continuum between traditional randomised RCTs and observational non-randomised RWD studies.[204] 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 subjects are recruited from practice settings. Randomisation is usually at the provider or clinical practice level and not at the individual patient level. Post-randomisation, patients and providers make treatment decisions and no specific efforts are put to assure patient adherence to the intervention (such as drug) 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 of patients with a condition for which there are 2 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 specialty 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 on Ethical and legal issues in using RWD).

3. Study Hypothesis, study treatment and comparator treatments(s): The primary hypothesis must be well-defined and 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 although the in real-world clinical practice may generally be clinical equipoise for the primary clinical outcome that was the basis for the RCT based regulatory initial approval. Treatment decisions follow routine practice as determined by the participating physician or participating 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 ER visits, hospitalisation, death etc.

5. Blinding: usually patients and physicians are unblinded to treatments. Outcome assessment and adjudication may be done in a blinded manner when it is possible to do so. 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 SAEs 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"[205] is a useful reference to use.

8. Statistical Analysis Plan: Design and statistical analysis approaches to address differences in baseline characteristics and impact of measured and unmeasured confounders will be dealt in other sections of the document. Needless to say, 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 comparison 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, co-interventions/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 conduct of the study.

b. Single arm trials using external RWD controls

External controls, typically derived from past traditional RCTs, have been used as a control arm for single arm trials. More recently, the use of 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/and where randomisation is highly unethical in context of a promising treatment for a serious disease with a high unmet need. Data from registries, administrative EMR/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[206] 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 drug is self-evident (e.g. general anaesthetics). ICH E10 (2001)[207] 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. Under its RWD framework program, US FDA is expected to issue guidance on use of non-randomised, single trials with external control derived from RWD.[208]

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 the 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.

Selection bias may be addressed to a certain extent by employing 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 preferably used.

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 the inferences from such studies.

Information bias may arise from differences in the collection, recall, recording and processing of information. When information bias is differential, it may result in exposure and outcome misclassification. 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 lead to an underestimate of treatment effect. On the other

hand, differential information bias could work in either way, resulting in an overestimate or under-estimate 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 the use of an appropriate definition of exposures and outcomes.

3.2.4 Bias and unmeasured confounding

When addressing the use of RWD it is important to realise that assessment of systematic error (bias) is a key element of any study that aims to evaluate a possible treatment effect of a medicinal product. However, we should acknowledge that the role of bias in (pharmaco)epidemiology has been described in many guidelines and reference works and therefore the aim of this paragraph should be to discuss to the most important forms 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.[209]

Unmeasured confounding

A distinction can be made 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. Therefore, in OSs, it is important to assess and question the potential impact of residual confounding. Because the impact on final results can be significant, it is strongly recommended to carry out sensitivity analyses. See section 3.3.7 on Principles of sensitivity analysis.

Selection bias

Selection bias relates to the selective recruitment of subjects in a study that 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.

- Referral bias

  Referral bias can occur if a patient is more likely to be recruited into a study due to this exposure status than a control patient with the same drug exposure status.[210] An example that has been referred to 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 that are easier to relocate.[211]

- Self-selection bias

  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.[212] 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

  A third example is prevalence bias 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.[213]

- Protopathic bias

  Finally, protopathic bias may relate to the issue of reverse causality. This can occur, for example, when a drug 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 drug and the cancer, even though the first symptom (headache) occurred before exposure to the drug. This is described in more detail by Jessica Chubak et al. [214]

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. in case controls studies cases and controls can have different recall of their past exposures) and surveillance or detection bias.

- Missing data

  "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 within patient registrations or health care databases than with RCTs or even carefully organised cohort studies, for several reasons. For example, it is unusual within registrations 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 registration (for example, if there is unequal registration).

- Surveillance bias

  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 an increased risk of bleeding that can trigger screening for endometrial cancers, leading to a higher probability of early-stage endometrial cancers being detected. 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.[215] This may also occur in a study in which a new treatment was assessed in a single arm trial and subsequently compared to historic controls (with no treatments).

- Immortal time bias

  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. [216] 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[217] 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".

- Other time-related bias

  Other forms of time-related bias. In many database studies, exposure status during hospitalisations is unknown. Exposure misclassification bias may occur with a direction depending on whether exposure to drugs prescribed preceding hospitalisations 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) hospitalisations that occur 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 use and multiple hospitalisations.[218]

3.3 Considerations for statistical analysis in a RWD setting

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.[219] 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 subjects who are at risk and not subjects without potential for experiencing the event at the time.[220]

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[221] 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 for 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 those 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). Studies have shown that communicating the magnitude of relative effects is improved when 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.[222] 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 a critical aspect of the analytic 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[223] and the Aalen-Johansen estimator[224] of the cumulative incidence of each event. Cumulative incidence probabilities can be estimated in consideration of competing risk events.[225] Group comparisons of the cumulative incidence probabilities over the whole time interval can be tested by using Gray's test.[226] 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 in the estimation of treatment effects adjusted or controlled for potential confounding variables. 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, that are 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.[227] 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 1 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 should be given when introducing a categorical variable with three or more

non-ordinal values into the model. Creating dummy variables can be introduced to such variables.[228]

Presentation of results: The presentation of the results of the analysis should not be misleading, thus needs to be well considered. In the simplest studies this presentation usually involves a table of risk ratios for the variables of interest with the 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.[229]

3.3.5 Time-dependent covariates and time-varying effects

Most of the variables discussed until the previous sections are known at the time when observation of the subjects begin, or "time origin". These are time-fixed covariates. Time-dependent or -varying covariates are those whose value may change after the subject entered the study. Examples include continuous variables such as WBC or neutrophil count after starting chemotherapy, or binary variables indicating whether the patient developed febrile neutropenia after initiation of therapy or whether the patient is discharged by a given time. Because the use of multivariable models to adjust for variables observed during follow-up can introduce bias, alternative methods based on weighting should be used.[230]

3.3.6 Matching approaches for comparators

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. There are two ways of matching: simple and propensity score matching.

In the 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.[231] 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 adjustment can be done using regression adjustment.[232,233,234,235]

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 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). Both 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.[236] Matching has some costs as well, including generally less precision due to exclusion of unmatched observations.

Propensity score analysis is a useful tool for causal inference in OSs, but it is important to note that it relies on several assumptions, including the correct specification of the propensity score model and the absence of unmeasured confounders. These assumptions are necessary for causal inference in general, but our confidence in being able to fit a well-specified model may be

greater for the relations between the covariates and the outcome, if there are sufficient number of outcomes to support such a model. Use of multiple analytic strategies as a sensitivity analysis (see the next section) or doubly-robust estimators[237] may serve as a useful approach, drawing strengths from both strategies.

3.3.7 Principles of sensitivity analysis

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 as repeatedly described. Sensitivity analysis is a series of analyses conducted with the intent to explore the robustness of inferences from the main estimator to deviations from its underlying modelling assumptions and limitations in the data, thus can help address these issues to different scenarios, assumptions, and sources of variability.[238]

Analyses results are considered to be "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 reproducibility of the analysis, increased confidence in the findings.

Traditional sensitivity analysis is to test basic assumptions such as variable definitions and to consider the impact of an unmeasured confounder. A study's underlying assumptions can be altered along a number of dimensions to evaluate robustness of results, including study definitions by modifying exposure/outcome/confounder definitions, 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.[239,240] Subpopulations such as paediatric-, geriatric-, 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. Sensitivity analysis results can be presented in tables or graphs. Tables should allow readers to determine the influence of changes in assumptions. Graphs are useful when the exposure and/or outcome variable being modelled is continuous. Of note, it is important to balance the benefits of including numerous sensitivity analysis results with the need for concise reporting.

3.3.8 Missing data

Incomplete data is a reality in all research but may be more extensive outside of the traditional randomised clinical trial. Missing data are defined as values that are not available and that would be meaningful for analysis if they were observed.[241] The extent to which data are missing and underlying dynamics that led to the missingness are important to consider when determining the approach to handling these in the analysis. None of these methods will entirely make up for lapses in data collection, but the negative impacts can be mitigated to some degree. In depth discussions of methods to address missing data are available elsewhere.[242] Recently, STRATOS (STRengthening Analytical Thinking for Observational Studies) initiative has published guidance framework for the treatment and reporting of missing data in OSs.[243]

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 that

is 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 BMI are missing for 10% of the study population, a predictive model would be fit among those with non-missing BMI data to estimate the likely value for BMI for those individuals where it is missing conditional on their age, sex, etc. In order to account for the uncertainty that is 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. In order 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 in conjunction 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[244] discusses the need for sensitivity analyses under the assumption that MAR is not valid.

3.4 Evidence-generation process, study registration, transparent reporting, audit trails and responsible communication

Taking regulatory decisions impacting public health in the form of MA approvals, and to some extent also reimbursement decisions, has traditionally been based on clinical trials for which rigorous criteria to ensure data integrity have been developed. This includes e.g. registration of protocols, pre-specifying analysis, blinding subjects, investigators, endpoint adjudicators and analysts, publication and results disclosure.

Similarly, the trust in RWE by regulatory bodies will be promoted 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 RWD studies would be conducted, it would encourage the 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 include 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, ISPE, and ISPOR have issued recommendations.

A joint task force of the ISPOR and the ISPE recommended that investigators pre-register their RWE studies and post their study protocols in a publicly available forum before starting studies to reduce publication bias and improve the transparency of research methods. Recognising that there are structural and practical challenges, the RWE Transparency Initiative has outlined a pathway how to move forward.[245]

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, meant to evaluate an effectiveness hypothesis. Thus, criteria for HETE are proposed to ensure specifically transparency and trust.[246]

The formulated general principles highlight the need to prospectively defining study methods in evidence generation, registration, stakeholder alignment with regulatory authorities/HTA 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 could therefore be viewed as general recommendations.

Box 1: ISPE/ISPOR taskforce recommendations for HETE

Source:[247]

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 prior to conducting the study analysis.
3. Publish HETE study results with attestation to conformance and/or deviation from the study protocol and original analysis plan. Possible publication sites include a medical journal, or a publicly available web-site.
4. Enable opportunities to replicate HETE studies (i.e. for other researchers to be able to reproduce the same findings using the same data set and analytical approach). The ISPE companion paper lists information that should be reported in order 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 work to publicly address methodological criticisms of their study once it is published.

7.	Include key stakeholders (patients, caregivers, clinicians, clinical administrators, HTA/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 registration of 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 drugs, biologicals and vaccines, including over-the-counter products following company internal standards and recommendations of trade organisations of pharmaceutical manufacturers. Also, Guidelines for Good Pharmacoepidemiology Practices (GPP) may apply.

In one company's example, these recommendations include registration of prospective OSs 10 days before study starts on ClinicalTrials.gov and company website located "Trial Finder"s.

Retrospective OS (secondary data collection OS) in patients focusing on the evaluation of efficacy and/or safety of an individual company drug are registered 10 days before study start on ClinicalTrials.gov and company website Trial Finder. Registration is currently not done for retrospective OS in patients not focusing on the evaluation of efficacy and/or safety of an individual company drug but the study results/outcomes are of significant medical relevance as assessed by a Bayer medical expert, excluding no-drug OS and disease OS. For all these types of OS, study result synopses are web posted 12 months after completion of the study, independent on the peer reviewed publication process.

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

Previous proposals have called for the registration of noninterventional studies[248,249,250] but the systems used and incentives to systematically register all studies have been unsatisfactory so far. It is hoped that 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 this will improve. 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 RWE 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 for pre-specifying 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 (PS) model are blinded to the outcomes.

It is noted by ISPE/ISPOR that in RWE, varied terminology is used around the world for the same concepts. 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.[251]

3.5 Reproducibility of RWD studies

Reproducibility is a cornerstone of the scientific method. However, there have been concerns about the reproducibility of research in many scientific fields, including cancer biology,[252] psychology,[253] and economics,[254] as well as clinical trials.[255] There have been several efforts to evaluate the replicability of studies in various disciplines, with the results suggesting that there is room for improvement. Recently, there was a systematic attempt to measure how we are doing in terms of the reproducibility of RWD studies.[256] This project identified a systematic random sample of RWD studies published in leading medical and epidemiology journals – 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 the majority of 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 methods, and to understand 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 make assessments about the validity and/or 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. 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 use of structured protocol templates like HARPER,[257] registering protocols, sharing code, and providing sufficient information on data sources.

3.6 Agreement between multiple RWD studies and RCTs

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

Bias is the issue that decision makers are most concerned about 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.[258] The credibility of RWD studies has suffered from this issue of 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.[259] Over 30 trials were systematically sampled from a variety of clinical areas and emulated using RWD. Some of the main take-aways from this project included:[260]

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 deeply 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 that is 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[261] and other similar RCT emulation projects (Observational Patient Evidence for Regulatory Approval and uNderstanding Disease (OPERAND),[262] Center of Excellence in Regulatory Science and Innovation (CERSI)[263]) 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.[264,265]

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 from the toolbox 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 drug effects under "ideal" conditions or clinical practice conditions?

3.7 Quality of RWD studies

Various tools exist to assess the quality of non-randomised studies such as STROBE (Strengthening the Reporting of Observational Studies in Epidemiology and GRADE (Grading of Recommendations, Assessment, Development and Evaluations). STROBE provides a checklist of

items that should be described in any reports of OSs.[266] 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. GRADE provides a transparent framework for developing and presenting summaries of evidence and provides a systematic approach for making clinical practice recommendations and has been officially endorsed by over 100 organisations worldwide.[267] GRADE has four levels of quality of evidence (very low, low, moderate, and high). 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.

References

[170] *A Framework for Regulatory Use of Real-World Evidence. September 13, 2017. Duke Margolis Center for Health Policy (Website accessed 2 June 2023)*

[171] *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*

[172] *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*

[173] *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. doi: 10.1093/aje/kwac140. PMID: 35916384; PMCID: PMC10144617.*

[174] *Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards. (Website accessed 4 June 2023)*

[175] *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.*

[176] *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).*

[177] *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*

[178] *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.*

[179] *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*

[180] *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*

[181] *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*

[182] *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*

[183] *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*

[184] *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*

[185] *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*

[186] *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*

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

[188] *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*

[189] 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

[190] 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

[191] 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

[192] 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

[193] 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

[194] 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

[195] 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

[196] 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

[197] 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

[198] 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

[199] 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

[200] 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

[201] 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

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

[203] 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

[204] PRECIS-2. (Website accessed October 2022).

[205] 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).

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

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

[208] 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. (PDF accessed 9 May 2023).

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

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

[211] Statology. What is Referral Bias? (Website accessed October 2022)

[212] 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.

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

[214] 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

[215] 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.

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

[217] 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.

[218] 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

[219] Pagano M, Gauvreau K. Principles of Biostatistics. 2nd edition. Pacific Grove, CA: Duxbury; 2000.

[220] 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

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

[222] 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

[223] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496-509. DOI: 10.1080/01621459.1999.10474144

[224] 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.

[225] 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 .

[226] 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.

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

[228] 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

[229] 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

[230] 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

[231] 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

[232] 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

[233] 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

[234]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

[235] 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.

[236] 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

[237] 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

[238] 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).

[239] 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

[240] *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.*

[241] *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*

[242] *Little RJ, Rubin DB. Statistical analysis with missing data. John Wiley & Sons; 2019.*

[243] *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*

[244] *Website access 4 June 2023*

[245] *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*

[246] *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*

[247] *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*

[248] *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*

[249] *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*

[250] *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*

[251] *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*

[252] 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

[253] *Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015;28;349(6251):aac4716. https://:doi:10.1126/science.aac4716*

[254] *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*

[255] *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*

[256] *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*

[257] *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*

[258] *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*

[259] *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*

[260] *Reference to be added. Publication in press.*

[261] *RCT DUPLICATE. (Website accessed 2 June 2023)*

[262] *Observational Patient Evidence for Regulatory Approval and uNderstanding Disease (OPERAND). (Website accessed 2 June 2023)*

[263] *Center of Excellence in Regulatory Science and Innovation (CERCI). (Website accessed 2 June 2023)*

[264] *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*

[265] *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*

[266] *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*

[267] *British Medical Journal (BMJ). BMJ Best Practice. What is Grade? (Website accessed 15 October 2022)*

2880

Chapter 4: Ethical and legal issues in using RWD

In the introduction of this report, two situations were indicated that point to the need for a change to the current "gold standard" of RCTs: the perceived requirements of efficacy tend to limit the participants in studies to a group that excludes vast numbers in the population, and more specifically excludes under-served populations such as patients from ethnic minorities, those from older age groups, and those with comorbidities; and, the reality of modern drug development is that that we need to start considering different data sources in addition to RCT data, to support our evidence generation for drug development. The introduction through to Chapter 3 of the report show strong evidence that there is 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 and normative level. Indeed, this has already begun, with some jurisdictions moving to include RWE in key statutes around the regulation of drug development.

RCTs may still remain an important part of the evidence base, but the scientific argument is clear that RWE must increasingly augment and replace evidence from RCTs to improve decision making. 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 being included, namely:

1. Given the shortcomings of RCTs, is it ethical to continue without integrating other forms of evidence?
2. What ethical and legal issues need to be taken into account 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 legal questions that should be addressed, 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 reuse of personal data in different settings from where it was initially gathered. There are also fundamental questions about how data and evidence about individuals in "real-world" settings are constructed, and how the observed-self relates to the legal, individual self, and the self in community. Within these broader questions, whose voices are heard 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. The current standards and expectations are built on a series of normative assumptions, and this chapter is designed to open those assumptions up for discussion - to create space in the normative discourse to explore the scientific proposals for change.

4.1 The current normative landscape

The current RCT-based regime, where trials are regarded as the gold standard of evidence, depends upon a number of normative presumptions, namely legal, ethical, and customary oughts. The key ought in the current regime is to compare, in real time, the effects of a new intervention/procedure against the effects of either no intervention/procedure or the current established equivalent, and that it is imperative to do so in an environment that removes biases caused by knowledge of which process is being used by each participant, and by underlying baseline factors such as comorbid conditions. This is predominantly a scientific question: is the presumption correct that this is the best way? However, there are normative questions that attach to this as well.

Legal presumptions - laws requiring particular conduct - flow from the science. While laws derive from societal values, there is no normative requirement that those laws take a particular shape or content beyond whether the science that they require and reflect is the best science. The Law is neutral as to its content in this regard. The chief imperative for the Law is that the science is good science - the best at any time. Therefore, that the Law currently requires a particular approach is not of itself a necessary condition for the future shape of the Law. Laws can change, and must change the face of improving science.

Ethics presumptions are different. Ethics requires particular behaviours - ethics concern what one ought to do at a deeper level than simply because the rule requires it, even by the consensus of democratic opinion. Ethics makes a fundamental appeal to the rightness of an action that transcends the particulars of the rule. However, ethics' weakness is that it is contested; there are competing claims to what constitute that ought. Further, ethics does not provide a necessary mechanism for the adjudication between competing oughts and constructions of those oughts.

There are a number of consensus positions, however, that dominate current bioethics: respect autonomy; do no harm; seek to create benefit; ensure justice.[268] How these basic principles apply in practice is a matter of debate, i.e. the granular realisation of the ideas is also contested, but a number of positions might be taken in relation to RCT.

1. Do no harm and seek to create benefit. The principle of avoiding harm to participants is paramount in RCT thinking and that is a route to achieving beneficence for the wider patient population. Therefore, identifying participants from whom the most robust results will be gained is imperative. Likewise, potential participants who are at risk of harm should be protected, and in the most part, that requires exclusion from the cohort.

2. Respect autonomy. This exclusionary principle comes, in no small part, from the operation of autonomy, or at least a protectionary operation of autonomy. The individual should not be exposed to risk of harm if possible; with the risk of harm extending to the foetus. This principle is interesting, as the protectionism overrules the individual's autonomy to choose to accept a risk to participate. This, perhaps, reflects the need for the reliability of the evidence, and public trust and confidence in the safety of trials.

3. Ensure justice. This limiting effect on the participants is a matter of justice. Two sorts of limitations operate: a first is about limits because of actual vulnerability (those with comorbidities, for example); a second, however, is a limitation through perception of vulnerability and reliability. Why is it that the profile of a perceived good candidate for a trial is a particular narrow profile? Is it a matter of actual or perceived vulnerability and reliability? The issue, then, is which populations are chosen to be engaged in trials, who are seen to be likely candidates to be in trials, whose voices and experiences are heard, and who is represented in RCTs. This is an ethical question, and the particular framing of the scientific requirements has produced, arguably, an ethically difficult result. The job is to ensure that, now that there is a strong scientific argument to change the paradigm to a methodology that encourages inclusion, law and ethics do not become barriers to that paradigm shift.

As a starting point it is worth considering whether informed consent is necessary as an ethics standard in data use in research. Clearly, RCTs work with this standard. RWE and RWD do not work easily with a presumption of informed consent as they depend on large, secondary use of already gathered data. Is informed consent necessary in all ethics theories? Considering this as a first case study helps to understand the contested nature of ethics.

Principlism, the use of established ethical principles to determine the right course of action, is well-established in bioethics. However, other theoretical ethics perspectives may be applied. Utilitarianism's imperative to act to maximise the utility for the maximum number of people could produce very different ethics requirements for trials.[269] It opens the door to a greater

expectation of participation: to requiring individuals to participate in trials. This would not be without precaution towards risk. Indeed, trials would have to be even more carefully considered and regulated for their potential harms to participants in order to ensure the utility of public confidence and trust in the pharmaceutical industry. However, there would be a greater sense of individuals being required to take the (regulated, mitigated) risk of participation for the ethical duty to participate to seek the greatest happiness of the greatest number.

More deontological positions, those based on rules, for example, those of Kant[270] or Rawls[271] , might at first thought be very restrictive, requiring high levels of autonomy and self-determination. This is the case when one considers how, for example, a trial sponsor should respect the autonomy of the individual (potential) participant. However, when considered from the perspective of the potential participant, the duties towards participation might be somewhat different. Indeed, there could be something of a duty towards solidarity from the deontological, liberal ethics. The Rawlsian perspective is, perhaps, easy to see. Rawls sees justice as realised through fairness. To enable this in decision making, he places the decision maker behind a "veil of ignorance" whereby one is ignorant of one's place in society upon return from behind the veil (thus disengaging self-interest from the decision-making process). From behind the veil, one would produce a decision that would be fair for the worst-off in society. Standing behind the veil of ignorance, one cannot rely on others to be the participants in trials if one hopes to benefit from the fruits of those trials - monetary payment is not enough. The worst-off member of society consideration would mean, of course, that a very vulnerable person would not require themself to participate in a trial risking personal jeopardy or jeopardy for the trial, but where one is able to participate, the imperative could be to do so. Likewise, an application of Kant's Categorical Imperative - to treat others as ends and not merely as means to one's ends - might require the same reasoning: I, in considering whether or not to participate in a trial, should see myself as compelled to participate if my not participating would result in treating another as a means to my ends. Again, personal and project jeopardy would preclude my participation, but where I can participate, I cannot instrumentalise others by a refusal to participate.

This goes to the heart of a difficult impasse in which we find ourselves: modern (Rand[272]) liberalism, that has become dominant since the 1980s seems to vindicate as ethical that the only duty I have towards participation in trials and the development of new treatments is purely through the purchase of the products when they arrive at the market. Ethics, in this iteration, provides no compulsion to participate in a trial. Since a trial is a matter of a physical intervention, with risk attached, perhaps this is justifiable. Is it acceptable to require another person, through law based on their personal ethical obligation, to assume a physical risk? However, does the same apply to the (often theoretical) risks that apply in the case of the secondary use of already gathered data? At the heart of RWD and RWE is the presumption that the data and evidence reflect the real world? Where the ethics presumption is that the individual can opt out, at what point is the "real world" no longer real?

4.2 Ethical arguments for incorporating more RWE

As mentioned in the introduction, two of the main reasons for ensuring that a sufficient ethical and legal framework exists for using more RWD and RWE are that the old, gold standard of RCTs relies on data gathered from a very small subset of the population, and, second, that such data are increasingly being used. These justifications will now be set out in more depth.

This move towards broader use of RWE to evaluate efficacy as well as safety is justified not only by a need for stronger 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 regarding efficacy from RCTs may not translate into evidence regarding effectiveness in clinical care. This is because the actual patient population is often not well represented by typical participants in RCTs, who are often younger and healthier than many patient groups treated in daily practice. Clinical trials also tend to under-report harm, further weakening the evidence base for real-world clinical care.[273]

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.[274] 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 typically relies on evidence from clinical trials that describe the likely benefits and toxicities. However, 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. Thus, generalisability to typical patient populations treated in daily practice is often limited. Kennedy-Martin et al 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.[275] 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 treated according to guidelines more often and received more in-hospital procedures. Strict selection criteria for RCTs meant that participants were at a much lower risk of adverse events compared with patients treated in clinical practice.

If the efficacy-effectiveness gap means that patients are being given inaccurate information about the potential benefits and risks of treatments, then decisions made using that information may be being made without valid informed consent, disrespecting patient's autonomy and putting them at risk of avoidable harm. The efficacy-effectiveness gap also raises important issues regarding justice; if resource allocation decisions, including which treatments are funded, are made using evidence that is biased by the efficacy-effectiveness gap, then those decisions will also be flawed, with potentially wide implications for patients. Ultimately, the efficacy-effectiveness gap undermines the gold standard status of RCTs, and actually suggests that it would be unethical to continue with such a flawed representation of real-world effects on patients. Increasing use of RWE is one important way to fill the efficacy-effectiveness gap and augment the evidence from RCTs. This should not be seen as dropping the gold standard, or diminishing the standard of evidence required; rather, enhancing and supplanting RCT evidence with RWD can instead be viewed as reinforcing the gold standard with platinum plating.

Next, we have the fact that RWE is increasingly used in practice, and this often takes place without any ethical or legal framework specific to use of RWD being in place, even if frameworks for clinical trials exist in all jurisdictions. Particularly in the context of the COVID-19 pandemic, personal data was used to inform decision making on a scale not seen before. As well as the examples provided in chapter 1 regarding SARS-CoV-2 drugs and vaccines, and resolution of uncertainties in a post-approval phase, Polymerase chain reaction (PCR) test results were used to inform public health authorities about trends in infection and transmission, and RWD from hospitals on COVID-19 hospitalisation and intensive care occupancy was also an essential source of information. Later, data on vaccination rates also played an important role in evaluating the public's level of protection against the virus. Much of the data used in this collective effort was anonymised when combined for public health purposes, but at the individual level, personal data including test results, vaccination status location and contacts with others was shared to facilitate the public health response, in some cases before any new framework was developed.

A more specific example concerns a COVID-19 clinical trial conducted in over 50 hospitals across seven provinces in Canada. Consent was obtained from patients to link occurrence of death data with intervention and administrative data at each hospital. Originally, there was an interest in linking with occurrence of death data at 12-months post intervention and pooling data centrally, but currently, only meta-analysis using aggregate data would be possible which would provide aggregate survival percentages (in any case comparison of outcomes by province was not central to this research).

The justification for the study was that individual-level data are necessary for analysis to inform clinical decision making and understand long-term outcomes. Fact-of-death was selected as the lowest hanging fruit variable in administrative data, and the focus was on testing the process for linking with administrative data in a repeatable, scalable way. As part of the project a normal policy analysis of data access process is underway, following the project in real-time across centres. This enables documentation of key activities, obstacles, enablers to data sharing for secondary use in research, ultimately informing data holders on barriers to data access, and providing leverage for change in policy and practice. Studies like this highlight both the pressing need for using RWD in the pandemic and medicine more widely, and the potential obstacles to doing so in terms of current/outdated ethical frameworks and legal restrictions on data sharing.

4.3 Potential ethical issues in using RWD

Before proceeding to consider privacy and data protection concerns regarding the use of RWD, we should note that relying more on RWD also carries its own potential disadvantages. While it is true that RCTs suffer from the aforementioned 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. 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 to inform decision making, the problems with RCTs would simply be replaced with a new set of problems, resulting in an equally flawed evidence base. It is outside the scope of this chapter to explore how this required reliability can be ensured, but as stated in chapter 1, it is likely that an evaluation of the methodology 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.

With the exception of 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, with some HCPs in the UK simply displaying posters informing patients about this. If RWD is to be used more, then routine data linkage with patient records for the purposes of R&D may be a next step, and it might be argued that seeking informed consent for such use is disproportionate. In pragmatic clinical trials and comparative effectiveness trials, it is already accepted that consent may not be necessary where randomisation is not taking place;[276] others have argued that randomisation alone should not be the decisive factor in determining whether consent is necessary. 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. However, any such change in paradigm cannot be accomplished by diktat; societal discussion would have to precede any such legislative change.

4.4 RWD, privacy and data protection

RWD concerns at least in part the secondary processing of already-gathered data. Whereas the gathering of data prospectively gives a chance to be able to determine better parameters for the use of those data, this presents a number of problems in Data Protection law internationally.

The current operation of the gold standard of anonymisation and informed consent has produced a situation that feels strangely anomalous. 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 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 data subject to protect their 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 law 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.

4.4.1 The broad data protection landscape, using the EU legislation as a case study.

From its common international roots in the late 1970s,[277] data protection law has shared a common language and basic shape.[278] 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 individual citizen themselves to whom the data relate (data subjects). Persons with duties can be both legal and natural persons. 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 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; to act with integrity towards the data.

Lawful processing is prescribed to include (although not exclusively) two fundamental elements: 1) processing must be on (at least) one of the a legal bases for the processing of personal data (and in the case of the processing "sensitive personal data" - which includes medical personal data - satisfying one of the specific legal bases for lifting the general ban on processing such 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 a wide range of 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.

4.4.2 Specific issues in data protection and RWD processing

Next we would like to consider the major unresolved conceptual and technical issues for the use of RWD. This is not to suggest that data protection is an inappropriate obstacle or barrier to processing; far from it. There is a very 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 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 effective. 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. Citizens at the same time hold aspirations and concerns about the development of new therapies to cure and prevent illness, and about their privacy and the use of their personal data in different contexts. Further, 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. The 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, there must be a management of competing expectations between all the parties. For example, individual data subjects cannot expect cutting-edge pharmaceutical product development but also a complete opting out from 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. RWD implies an altruistic society that must be realised through its regulatory and governance structures.

4.4.3 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 the new situation. And they are gathered from many sources to create the image of the real world. 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 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. Personal data are also gathered, in many jurisdictions and 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 relating to 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 one purpose. Therefore, the first question relating to the processing of RWD is: does the original legal basis for processing cover this 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 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.

4.4.4 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 are issues that 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.

4.4.5 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 Research Ethics Committee 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. A distinction is made between direct and indirect gathering of personal data, that when data are gathered directly from a data subject the information must be provided, whereas where 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. 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. 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 in the case of 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. Costly here is interesting. On the one hand, there is a simple economic cost that might not be affordable in a

research project. However, the requirement here is more that one must take into account the potential damage to the data subject.

4.4.6 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, address, etc.) replaced with a code. The effect of this 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 is capable of re-identifying the individuals. Data protection law sees pseudonymised data - both the code and the coded data in our 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 are 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 have been made sufficiently to render the data as de-identified and no longer personal in the hands of the researcher. In other jurisdictions a harder line is taken, whereby the very possibility of the reconnection of 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 a pragmatism in the letter of the law, but requires harmonisation in the interpretation of the idea to ensure consistency.

One aspect that is interesting 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 a number of data controllers and those (no-longer) personal data are then passed to researchers. This would, in the above scenario, in many jurisdictions remove the data from the scope of the data protection law. However, the data remain identifiable in the hands of the data controllers. The researchers then run into a question about their dataset and send a question to the data controllers from whom they receive the data, which is answered 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 work arounds are tried against the backdrop of differing approaches by Data Protection Officers and Data Stewards, but RECs and IRBs (Institutional Review Board), and, too often, with very little guidance from the regulators before intervention for breaches are made.

4.4.7 Research or safety evaluation

In clinical trials and drug production, it is very interesting to observe that much of what has been discussed above in relation to research does not apply to the conduct of evaluations for safety of drugs 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.

4.4.8 Other jurisdictions

While the specifics of data law will of course vary between jurisdictions, many other countries adopt an approach somewhat similar to that of the EU with 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 drugs, vaccines, and other health-related supplies reach the market."[279] By enabling regulatory harmonisation, the AMA and its associated treaties will also facilitate the use of RWD and RWE.

In Brazil, the General Law for the Protection of Personal Data (Lei Geral de Proteção de Dados Pessoais, or LGPD) features similar key principles of data processing and privacy by design. 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 and purposes."[280] 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."[281]

In Canada, the federal Personal Information Protection and Electronic Documents Act (PIPEDA) has governed data use for over two decades, but in addition, each different province has its own health privacy law. Examination of these specific laws is outside the scope of this chapter, but PIPEDA diverges from GDPR in a number of 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 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 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. GDPR is also stricter in terms of data impact assessments, making them mandatory in certain circumstances, while PIPEDA only recommends them. Finally, "the GDPR and PIPEDA are also 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 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 institutional ethics committee approval. Such approval is also required for sharing between institutions.

4.5 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 its focus on perfect patients who are often highly unrepresentative of the populations in whom new drugs and other interventions will be used, combined with almost complete 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.

Given that many treatments are currently prescribed based on old and unrepresentative RCT evidence, this means both that patients may be being prescribed drugs 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 and nonmaleficence are both threatened by us 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, use of unrepresentative data across health systems threatens the principle of justice by distributing resources according to similarly flawed decisions. Equally, of course, any RWD used 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.

In turn, if it is vital to use more RWE, this means that 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 via ensuring informed decision making about medicines, rather than traditional protection of autonomy by keeping personal data siloed and sealed off, are likely to be paramount in increasing utilisation of RWE in an ethically robust manner.

4.5.1 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. 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, these all 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 RWCs, if not under their observation and monitoring. Again, the purpose of the data protection legislation is to safeguard the interests of the data subjects. What is crucial is that the potential abuse of those citizens through the misuse of their personal data be properly evaluated

and then avoided through robust and effective safeguards. What is inexcusable is that ineffective and outdated measures are used that enable personal data to be processed without proper regard to the dignity of the data subjects, whilst at the same time creating barriers through the inappropriate nature of those old concepts to legitimate data processing for ends desired by ordinary citizens that are equally protecting of their interests.

Perhaps COVID-19 is a beginning to a change in the approach. It is increasingly said that the pandemic brought an alignment of incentives in relation to 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 vaccines to respond to it. RWD came to the fore, and the pre-pandemic paramountcy 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 sky did not fall in. Almost in the same breath, the reversal of Roe v. Wade in the United States Supreme Court has dealt a massive blow to individual privacy. This is not only at the decisional privacy question of who decides, the State or the woman, but at the informational privacy level of how will, for example, information about menstrual cycles generated by apps be used in possible criminal trials. It exposes how commercial sale of sensitive data, for example, purchasing a pregnancy test, can lead to targeted marketing of pregnancy and new-born care products, leading to potential abuse of women in violent and abusive homes or before hostile laws.

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 that produce harm. Commercial interests cannot be tone deaf to the context within which they seek access to individuals' data; individual citizens equally cannot be tone deaf to the competing claims they make on society. If commercial interests request altruism from their data subjects, they must respond in altruistic access to their products and the research; if citizens want the benefit of new therapies and pharmaceuticals they must acknowledge that this requires access to their data.

This returns us to the questions of ethics that started this chapter. The ideas presented in the last paragraph indicate the need for a discussion about the nature of our social contract. To answer the regulatory, normative and governance questions posed by RWD, we cannot rely on the current political approach that avoids hard moral questions. The decisions upon which the reimagining of data protection governance for RWD can be made in a piecemeal way with different jurisdictions relying on somewhat unstable work-arounds to muddle through. However, that does not create the robust environment that our desire for co-produced, democratic science demands. Only by opening the debate to explore the competing interests of all stakeholders and respecting the concerns and hopes of all parties, at an international level and without any prejudice in favour of the economically rich countries and individuals, can the environment that RWD requires be created. Ironically, the solution is available in plain sight in the current legislation; it is within our grasp. What seems beyond our reach is the will to ask the most important questions. What responsibility do I have to others? What responsibility do I have to producing robust, honest science? What responsibility do I have to ensure access to healthcare products as a part of the right to healthcare? What is my commercial responsibility in that regard? What is my responsibility as a patient and as a member of the public in that regard? What duty of confidence do I owe to anyone whose data I process? What can I demand about my data? Can I really demand absolute privacy?

4.5.2 Beginning to change the landscape

The answer to the last question is, "of course not!" Privacy is not an absolute right, it is held in balance with the rights of others in society. However, individuals have rights to dignity, and those must be negotiated by all stakeholders. As indicated above, there are routes through the legislation that can better facilitate RWD processing: using the public interest as a legal basis, clarifying expectations around compatible processing, de- and re-identification of personal data,

and the like. However, to end this chapter, two ideas could be explored to spark the public discussion of how we want our personal data to be governed in the biomedical arena: to whom do data belong, and is privacy the right conceptual starting point?

Taking the first question about interests in personal data, there is a very interesting difference in the rhetorical and colloquial language of data privacy and the legal rights to personal data protection. Even in the presentation of the GDPR, the language is strongly that the data subject owns the data in question. It is my data. However, the law is based on a human right to protection rather than ownership. Duties are created around the processing of data that relate to an individual. There is a large difference between the two. This could well be a Lockean distinction[282]: that one gains personal property through the added value brought to raw materials; or it could be grounded in the reluctance shared in many jurisdictions to give legal ownership either over parts of the body (and personal data is being seen as an extension of this, see the Declaration of Helsinki); or a reluctance to acknowledge ownership in information generally. Whatever the reason, the ownership of personal data is obscure. And this, in the context of medical information, is accurate. Who owns the data? If one gives blood at a hospital, there is an argument that the blood is owned by the donor (already an interesting property word denoting a transfer of title), and one could by extension say that the chemistry of the blood is owned by the individual. But the action to transpose the data stored in the raw material is that of the hospital through the operation of processing of the blood to separate the personal data from the physical chemistry. When that information, that blood, is processed by the researcher, and a new understanding is created from that novel processing (perhaps resulting in a patentable product), it is the work that generates the property, not the origin - the donor, again that work, giving up their claim like the seam of coal yielding to the miner's axe. But, in the age of bitcoin, could a new model allow a direct payment, perhaps cents, to the donor as the original owner of the data that is mined? Would that be appropriate? Would such a commercial contract strengthen or weaken our social contract? In particular, would it enable a global justice to prevail, or would it further strengthen institutional and social discriminations? This, as a first question, is very interesting.

The second question is the following: is privacy the most appropriate conceptual basis for data protection? Data protection emerges as a separate human right from the right to privacy (to a private life). From the earliest writings on the subject, privacy has tended towards an exclusionary right. This is not exclusively the case, neither is it necessarily the case, but it is a dominant conceptual flavour in data protection. For one different example, see Graeme Laurie's approach to privacy as a space where relationships between individuals can be negotiated rather than a presumed set of values.[283] He points also to a small number of cases in the The European Court of Human Rights (ECtHR) that open the idea of privacy being concerned with human flourishing. This is strange, as the earlier concept in relation to the processing of (medical) data was confidentiality. Confidentiality conceptually offers the negotiated terms by which information can be used for specified purposes. This is the purpose of data protection legislation. It is not designed to shut down or prohibit the processing of personal data, but rather to regulate it in such a way as to create an appropriate balance of safeguards for the processing of personal data for different, legitimate ends. Confidentiality has strong links to professional duty, to the duty to place one's clients' interests before one's own in acting in a professional capacity. A shift away from a privacy debate to a confidentiality debate offers an opportunity to re-focus the discussion, back to the starting point of asking how to enable data to be processed for legitimate ends and how to safeguard legitimate interests. The professionalisation of researchers, as is perhaps emerging in the drive to address research integrity, cannot come too soon to assist in this re-evaluation of what data protection is seeking to achieve, particularly in terms of using RWD.

References

[268] *Beauchamp TL, Childress JF. Principles of biomedical ethics. 9th ed. Oxford University Press; 2019.*

[269] *Mill JS. (Crisp R, ed.) (1998) Utilitarianism. Oxford University Press; 1998.*

[270] *For a good introduction to Kantian ethics, see Johnson R and Cureton A. Kant's Moral Philosophy, The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Zalta EN & Nodelman U (eds.), (Website accessed 15 April 2023).*

[271] *Rawls J. Justice as Fairness: a restatement. Harvard University Press; 2001.*

[272] *Badhwar NK, Long RT. Ayn Rand. The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), Zalta EN (ed.) (Website accessed 15 April 2023).*

[273] *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*

[274] *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).*

[275] *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*

[276] *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).*

[277] *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).*

[278] *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.*

[279] *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).*

[280] *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*

[281] *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*

[282] *Uzgalis W. John Locke. The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Zalta EN, Nodelman UI (eds.) (Website accessed 15 April 2023).*

[283] *Laurie G. (2002) Genetic Information: A Challenge to Medico-Legal Norms. Cambridge University Press; 2002.*

3534

Chapter 5: Conclusions and future directions

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

Though differences in engagement and process of submission by countries exist, acceptance of RWE in diverse phases of the product lifecycle has been observed 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/expectations are high quality data/information and reliability, accessing and understanding the information. Continuing the ongoing 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 RWE, as well as the generated RWE. These activities will strengthen the grounds for RWE acceptance and will support the development of evolving technologies and methods, including artificial intelligence and also open potentially the access to different sources of data, e.g. health care sensor for remote monitoring.

RWE could and should be considered, if appropriate, because strategy for addressing evidence gaps should cover all types of evidence generation, whether this includes 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 the effectiveness as well. While the common misunderstanding is that RWD includes only EHR, the scope is much broader. It also includes other sources such as SRSs and surveys.

Each RWD source has its strengths and limitations, and it may be useful for certain safety and effectiveness purposes, but not for others. Survey data sources are very useful to estimate the burdens of diseases, but they are not the most appropriate associations between medicinal products and outcomes, which require follow-up information. A scientific evaluation of the fitness of a RWD source for the purpose of the study is essential in choosing a data source.

The rapid development in new technologies has resulted in new RWD sources with large volumes extremely quickly. Although the current use of these emerging sources is still 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 a 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.

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 for inclusion. 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 the issue of 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.

There is a compelling scientific case for an extended utilisation of RWD, including data addressing specifically efficacy/effectiveness and this includes change at the regulatory and normative level.

In terms of the ethical framework, 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 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 in order to create space in the normative discourse to explore the scientific proposals for change.

The evidence-efficacy gap undermines the gold standard status of RCTs, and suggests that it would be unethical to continue with such a flawed representation of real-world effects on patients. Increasing use of RWE is one important way to fill the efficacy-effectiveness gap and augment the evidence from RCTs.

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

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 personal data protection regime is essential to the acceptance and operation of RWD processing.

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

The shape and structure of such frameworks will have to be discussed at the societal level, along with consideration of whether privacy is the most appropriate conceptual basis for data protection.

Using RWE to augment RCT evidence is an ethical imperative.

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.

To be effective and to foster trust and confidence, the data protection regime must equally be coherent, appropriate and effective. There is a strong argument that the regulatory regime should 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.

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 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.

APPENDIX 1: Case studies

These case studies complement the chapters in this report; they are not intended in lieu of guidance. We encourage all readers to follow local guiding principles and regulatory guidance 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 utilisation of natural history of disease data 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. What was the research question?	Do patients treated with fosdenopterin show an improved survival outcome compared to untreated patients in a natural history disease study?
Medicinal product.	Nulibry (fosdenopterin), a synthetic cPMP, was approved in 2021 by US FDA for treatment for Molybdenum cofactor deficiency (MoCD) Type A. There was no pharmaceutical treatment approved before fosdenopterin for this rare and fatal disease.
Indication/Disease treated.	MoCD Type A is a rare, neurodegenerative, autosomal recessive disease with an estimated U.S. 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 the pre-marketing clinical research.
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 Nulibry 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, noninterventional 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 nature of rareness 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, noninterventional, 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? What were the advantages? Disadvantages?	Data analysis used the log-rank test to compare treated and natural history control patients, and Kaplan–Meier (KM) plots and methods to estimate survival parameters for each group. Additionally, the SAP 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 Nulibry 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 Nulibry provides a survival benefit in patients with MoCD Type A.
Conclusion. Do you have recommendations or key learnings to share?	When designed and conducted properly, external controls from real world data sources can provide RWE in 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 provides assurance. The benefits of Nulibry outweigh its risks when used according to the product labelling.
Contact details.	Jie Li Jie.j.li@fda.hhs.gov

3648

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 the use of 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 common data model that provides a standardised way to represent and organise observational healthcare data from disparate sources, enabling data harmonisation and facilitating data sharing and collaboration across different healthcare systems and research institutions (see Chapter 1 on Uses of real-world evidence for decision making during the product lifecycle).
Study question. What was the research question in the example?	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, BNT162b2, mRNA-1273 and Ad26.COV2.S. The ChAdOx1-S and the Ad26.COV2.S vaccines use a weakened version of a common cold adenovirus. The adenovirus is modified to carry the genetic code for the spike protein found on the surface of 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 based on mRNA that codes for the spike protein found on the surface of the SARS-CoV-2 virus, which causes COVID-19. When the vaccine is given, the mRNA enters cells in the body and instructs them to produce the spike protein.
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.
Stage of the medicinal product development lifecycle.	The study took place at post-market stage of the vaccines analysed.

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 compared with mRNA-based COVID-19 vaccines, 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.
What were the data analysis methods used? Why were they chosen? What were the advantages? Disadvantages?	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. 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. 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. The use of 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 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 the 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.

Conclusion. Do you have recommendations or key learnings to share?	This study provides a key context on the complications in unvaccinated subjects suffering from COVID-19, showing these patients a remarkable increase in the risk of some outcomes, 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.
Contact details.	The study was published in the British Medical Journal in 2022. Reference: Li X, Burn E, Duarte-Salles T, Yin C, Reich C, Delmestri A 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 doi:10.1136/bmj-2022-071594. Available at: https://www.bmj.com/content/379/bmj-2022-071594 Corresponding author: E Burn Edward.burn@ndorms.ox.ac.uk Contact: Miguel A. Mayer, Hospital del Mar in Barcelona (Spain). email: mmayer@psmar.cat

3654

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 observed 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 COVID-19 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 on Real-world data sources).
Study question. What was the research question in the example?	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 post-COVID-19 disease. AESI are specific and significant adverse events that are potentially related to a particular medical intervention, such as a medication or vaccine. The identification and monitoring of AESIs is an important part of drug safety surveillance and regulatory decision making. By closely monitoring adverse events of special interest, regulatory authorities can ensure that medical interventions are safe and effective for patients.
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 can also occur. Disease severity varies by age and underlying health conditions. As of April 2023, the number of confirmed COVID-19 cases worldwide has exceeded 555 million, resulting in over 8 million deaths. On the other hand, 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: Guillen-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.
Stage of the medicinal product development lifecycle.	The study took place at post-market 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.
How did the involvement of RWD / RWE in the study affect the study design at the outset?	The number of patients included in the study using 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 bioinformatic tools. The total number of participants included in all databases was 945,520,607. OSs have been conducted to investigate the incidence rates of AESI among patients with COVID-19 and those who have been vaccinated against COVID-19. To accurately assess the risk-benefit 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 healthcare 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? What were the advantages? Disadvantages?	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.
What legal data protection requirements had to be met in the countries you were working in?	Informed consent from of individual patients was not required as anonymised information was obtained from the different clinical databases. The study protocol was approved by the different 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). Utilising 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. Do you have recommendations or key learnings to share?	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.
Contact details.	The study was published in the eClinicalMedicine journal in 2023. Reference: Voss E, Shoabi A, Yin Hui Lai L, Blacketer C, Alshammari T, Makadia R et al. Contextualising adverse events of special interest to characterise the baseline indicence rates in 24 million patients with COVID-19 across 26 databases: a multinational retrospective cohort study. eClinicalMedicine 2023;58:101932. Available at: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00109-8/fulltext Corresponding author: Erica A. Voss, email: evoss3@its.jnj.com Contact of case study: Miguel A. Mayer, Hospital del Mar in Barcelona (Spain). email: mmayer@psmar.cat

3660

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 NDMA-contaminated valsartan drug products by regulatory authorities worldwide was necessary in order to protect public health. The detection of different nitrosamine impurities in drug products since 2018 led to the introduction of a new threshold by the EMA.
Study question. What was the research question in the example?	Is there an association between filled prescriptions of potentially NDMA-contaminated valsartan drug products 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, N-nitrosodimethylamine (NDMA) was detected in the valsartan active substance manufactured by Zhejiang Pharmaceuticals. 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 inhibitor drugs, which block the conversion of angiotensin I to angiotensin II.
Stage of the medicinal product development lifecycle.	Post-market
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.
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, the AOK. On average, nearly 25 million persons were insured by the AOK each year during the study period. Furthermore, MAHs provided batch-related data on all valsartan drug products for the study period. This included information on which batches were manufactured using the potentially contaminated active ingredient valsartan and how many packages of these drug products were sold. The long time period (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? What were the advantages? Disadvantages?	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 routine data used for the study cannot be shared with or transmitted to third parties due to legal restrictions.

the countries you were working in?	
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 drug products. The study is an example of how to use pharmacoepidemiological methods and RWD to examine urgent questions of drug safety.
Conclusion. Do you have recommendations or key learnings to share?	The conclusion of the study included that careful monitoring of potential further effects of NDMA-contaminated valsartan after longer periods is advisable.
Contact details.	Prof. Dr. Britta Haenisch, britta.haenisch@bfarm.de, head of research division at BfArM, Bonn, Germany

3664

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

Topic	Summary Information
Rationale.	The risk of cardiovascular death in patients with gout was higher in the febuxostat group than in the allopurinol group 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 remains 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 1.7.3 on General RWE landscapes in various countries - Japan.
Study question. What was the research question in the example?	Is the risk of cardiovascular events associated with urate-lowering drugs 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. Benzbromzrone promotes uric acid excretion by inhibiting uric acid reabsorption in the tubules. Nonproprietary name: Febuxostat Branded name: Feburic Tablets MAH: Teijin Pharma Limited
Indication/Disease treated.	Febuxostat Indication:(1) Gout, hyperuricemia, (2) Hyperuricemia associated with chemotherapy.
Stage of the medicinal product development lifecycle.	Post-market
Where were the study protocols registered?	The protocols were registered with 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 the risk among different drugs.
What were the data sources used and why were they chosen?	Data from the National Database of Health Insurance Claims and Specific Health Check-ups of Japan (NDB) were used for analysis in this study, because (1) NDB is the largest database managed by the MHLW, collecting information on nation-based medical claims from hospitals, clinics, pharmacies, and dental clinics in Japan; and (2) the

	long follow-up period from hospitals where patients underwent treatment can be ensured.
What were the data analysis methods used? Why were they chosen? What were the advantages? Disadvantages?	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 contained in NDB is anonymised for protecting personal information, and does not include a 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 required 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. Do you have recommendations or key learnings to share?	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.
Contact details.	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. DOI: 10.1111/cts.13439 Yoshiaki Uyama, Office of Medical Informatics and Epidemiology, PMDA, Shin-Kasumigaseki Building, 3-3-2 Kasumigaseki, Chiyodaku, Tokyo 100-0013, Japan. Email: uyama-yoshiaki@pmda.go.jp

3669

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

Topic	Summary Information
Rationale.	To investigate the association between human granulocyte colony-stimulating factors (G-CSF) preparations (filgrastim, nartograstim, lenograstim, and pegfilgrastim) available in Japan and thrombocytopenia in patients treated with antineoplastic agents, a nested case-control study was conducted using the Medical Information Database NETwork (MID-NET®) with the cohort of the Japanese population taking antineoplastic agents. MID-NET® stores electronic medical records, administrative claim data, and diagnosis procedure combination data of about 5.3 million patients (as of December 2020) in cooperation with 10 healthcare organisations, including 23 university hospitals or regional core hospitals. See section 1.7.3 on General RWE landscapes in various countries - Japan.
Study question. What was the research question in the example?	Do G-CSF preparations cause thrombocytopenia in patients treated with antineoplastic agents?
Medicinal product.	G-CSF preparations (filgrastim, nartograstim, lenograstim, and pegfilgrastim) are human granulocyte colony-stimulating factors. Nonproprietary name: Pegfilgrastim (genetical recombination) Branded name: G-LASTA Subcutaneous Injection
Indication/Disease treated.	Pegfilgrastim (genetical recombination) Indication: prophylaxis of neutropenia caused by antineoplastic agents
Where were the study protocols registered?	The protocols were registered with PMDA.
Stage of the medicinal product development lifecycle.	Post-market
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 us 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 (occurrence of thrombocytopenia after administration of G-CSF preparations during the treatment period with antineoplastic agents) can be obtained in the same hospital, even though MID-NET® can only follow-up a patient within a hospital.
What were the data analysis methods used? Why were they chosen? What were the advantages? Disadvantages?	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 length, commodity, and co-prescribed drugs. 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 the occurrence of radiological therapy. Similar analysis was conducted on each drug in the detailed analysis.
What legal data protection requirements had to be met in the countries you were working in?	The data contained 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 G-CSF-induced thrombocytopenia.
Conclusion Do you have recommendations or key learnings to share?	A significantly increased risk of thrombocytopenia associated with pegfilgrastim was identified (aOR: 7.4 95% CI: 2.0–28.1). More attention on 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.
Contact details.	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(2): 473-479. DOI: 10.1002/cpt.2263 Yoshiaki Uyama, Office of Medical Informatics and Epidemiology, PMDA, Shin-Kasumigaseki Building, 3-3-2 Kasumigaseki, Chiyodaku, Tokyo 100-0013, Japan. Email: uyama-yoshiaki@pmda.go.jp

3674

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 Annex) 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.

Annex

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 - HSA Jun Zhao - NMPA/CDE Peter Mol - CBG-MEB
Federated and other Data Networks	Exploration of existing federated data networks used worldwide including their challenges and opportunities	***Melissa Kampman – HC*** Jesper Kjaer - DKMA Azumi Takano - PMDA Patricia Bright - US 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

(to follow)

APPENDIX 4: List of commentators (following public consultation)

(to follow)