Real-World Data (RWD) & Real-World Evidence (RWE)

Introduction

RWD and **RWE** are defined by the European Medicines Agency (EMA)[2] as a regulatory authority as such:

RWD: "routinely collected data relating to a patient's health status or the delivery of healthcare from a variety of sources other than traditional clinical trials."

RWE: "the information derived from the analysis of RWD".

Real-World Data (RWD) and subsequently Real-World Evidence (RWE) are derived by big data. **'Big data'** is defined as 'extremely large datasets which may be complex, multi-dimensional, unstructured and heterogeneous, which are accumulating rapidly and which may be analysed computationally to reveal patterns, trends, and associations[5]. Medicines regulators will increasingly use insights derived from big data to assess the benefit–risk of medicines across their lifecycle[47]. The HMA-EMA Catalogues serve as repositories of metadata from RWD sources and studies and can support regulators, pharmaceutical companies, and researchers in evaluating the use, safety, and effectiveness of medicines[32].

RWD studies and RWE studies each follow specific study formats tailored to their purpose. RWD studies primarily focus on collecting data from real-world sources, are typically observational and descriptive in nature, and are designed to gather information about patient populations, treatment patterns, or healthcare delivery. On the other hand, RWE studies aim at testing hypotheses or deriving conclusions from

the collected data using predefined methodologies to generate actionable evidence[2].

If you want to learn more about big data and digital health, visit EUPATI Open Classroom.

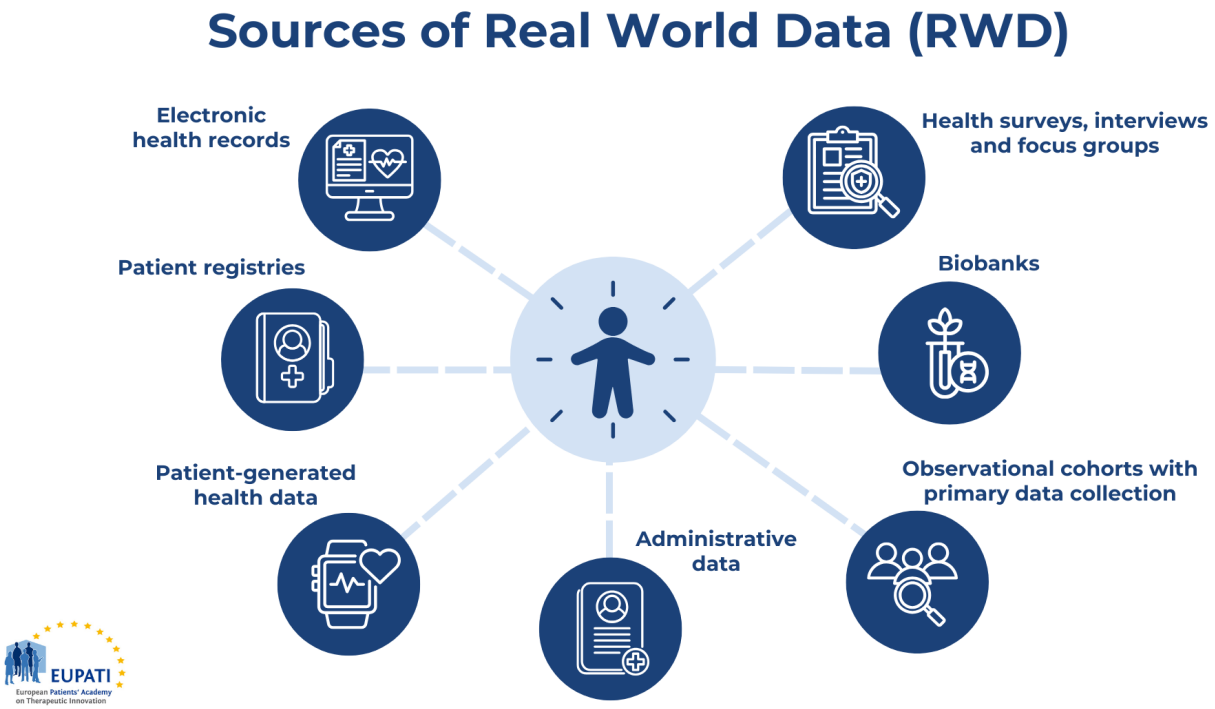

Figure 1: Sources of Real-World Data (RWD) – Created by EUPATI

Figure 1 exemplary shows primary sources of real-world data, which include electronic health records, patient registries, patient-generated health data, administrative data, observational cohorts with primary data collection, biobanks, health surveys, interviews and focus groups.

In the past, RWE has been accepted for post-approval safety monitoring but is now gaining recognition for evaluating drug effectiveness. Unlike traditional clinical trial data, which is collected under highly controlled conditions to ensure internal validity, RWE uses RWD to demonstrate the actual experiences of patients in everyday environments, capturing a broader range of patient behaviors, treatment responses, and health outcomes.

RWE has proven valuable for approval and reimbursement, and also for monitoring the safety of marketed products. Regulators such as the US FDA have leveraged digital systems, like the Sentinel Initiative, to track post-marketing safety outcomes[20]. Moreover, RWE supports healthcare providers in making informed decisions, with the UK National Health Service (NHS) using it to negotiate cost rebates for specific treatments. Payers are increasingly using claims data and health technology assessments (HTA), such as those by the UK's NICE, to make data-driven pricing and reimbursement decisions[9].

In recent years, both regulators and payers have increasingly integrated RWE into their decision-making processes, as demonstrated by the EMA initiative to establish the Data Analysis and Real-World Interrogation Network of the European Union (DARWIN EU)[1]. Pharmaceutical companies are also rapidly expanding their use of RWE to inform trial designs, improve clinical guidelines, facilitate reimbursement discussions and support the market access of new medicines.

Various initiatives have been launched to harness the insights embedded in RWD. For instance, the Innovative Medicines Initiative's (IMI) Big Data for Better Outcomes (BD4BO) program and the Horizon Europe framework have fostered numerous projects, including EHDEN, PIONEER, HARMONY, IDEA4RC, and the Data Saves Lives Initiative each dedicated to advancing RWD analysis and utilization. Additionally, the GetReal Institute, developed from an earlier IMI project, is working to enhance the credibility of RWE by establishing best practices for RWD collection and analysis while promoting collaboration among key stakeholders[33].

Real-World Data (RWD)

Table 1 shows that RWD refers to data on patient health or experience collected outside of highly controlled clinical trials. RWD can come from different sources:

Types of RWD	Descriptions
Electronic health records	• Computerised individual patient records. • These are typically used to inform the clinical management of patients. • These may also integrate data from other information systems including laboratory, genomic, and imaging systems.
Patient registries	• Registries are organised systems that collect uniform data (clinical and other) to identify specified outcomes for a population defined by a particular disease, condition, or exposure. • Registries can serve several purposes including research, clinical care or policy. • Registries can include interventional studies.
Patient-generated health data	• Data generated directly by patients or their carers including from wearable, medical or personal devices, mobile apps, social media, and other internet-based tools. • Data can be collected actively (for example, by people entering data on a form) or passively (for example, a smartwatch that measures people's activity level, or remote monitoring devices such as ECG devices, respiratory monitors etc).

Administrative data	• Data collected for administrative purposes by health and social care services. • Dataset contains information on diagnoses and procedures done for patients. They do not contain information on outcomes or effectiveness.
Observational cohorts with primary data collection	• Traditional prospective studies designed to answer one or more research questions. • They may link to patient records for health outcomes.
Biobanks	• Biobanks store biological samples and their related data on individual persons. • Biobanks can be collected and used for clinical practice and research. • Samples and related information can include demographic, diagnosis, medication, genomic, metabolic, laboratory, imaging, and lifestyle data.
Health surveys, interviews and focus groups	• Health surveys involve the systematic collection of data about health and disease in a human population through surveys. • They have various purposes including understanding trends in health in a population or understanding patients' experiences of care. • Interviews and focus groups are done to collect qualitative data such as patient perceptions and experiences.

Table 1: Types of Real-World Data – Created by EUPATI

RWD can be an alternative and/or supplement[1] to randomised controlled trials (RCT). Randomised trials may not be available for different reasons, including in cases where:

- randomisation is considered unethical, for instance, because of high unmet needs,
- patients are unwilling to be allocated to one of the

interventions in the trial,
- healthcare professionals are unwilling to randomise patients to an intervention they consider less effective,
- a small number of eligible patients,
- financial or technical constraints on studies,
- not all treatment combinations (including treatment sequences) can be directly assessed.

Randomised controlled trials may be complicated to perform for rare diseases, innovative and complex health technologies, or in certain populations. Similarly, randomised controlled trials can be challenging for medical devices and interventional procedures for different reasons:

- difficulty of blinding,
- the importance of learning effects,
- changes to standard of care making the choice of comparator challenging,
- changes to the characteristics of the technology over time that may impact on performance,
- limited research capacity or access to funding[3].

Furthermore, RWD can provide information on disease epidemiology, causal estimations, actual treatment plans, treatment practices, quality, effectiveness, and patient outcomes.

[glossary_exclude]RWD in Health Technology Assessment (HTA)[/glossary_exclude]

RWD can also be used in Joint Clinical Assessment (JCA) to define a relevant population, identify comparators, and demonstrate safety, (comparative) effectiveness and outcomes.

These aspects are typically part of the PICO framework[41], which is integral to defining research questions.

[glossary_exclude]To learn more about the PICO Framework, explore the relevant course in EUPATI Open Classroom and visit the Mini-course Starter Kit on the Health Technology Assessment Regulation (HTAR) in EUPATI Toolbox.[/glossary_exclude]

Real World Evidence – RWE

RWE is based on RWD. When RWD is analysed to answer scientific questions – thus use to generate evidence – it is called Real World Evidence (RWE).

RWE refers to insights gained from data collected outside of traditional clinical trials, such as electronic health records, registries, wearables, and medical devices. Unlike clinical trial data, RWE data is initially gathered for other uses and then repurposed for research. This diversity allows for a wide range of analyses, from patient and condition profiling to economic modelling and comparative effectiveness studies. The choice of study design in RWE depends on the specific research question and its intended application.

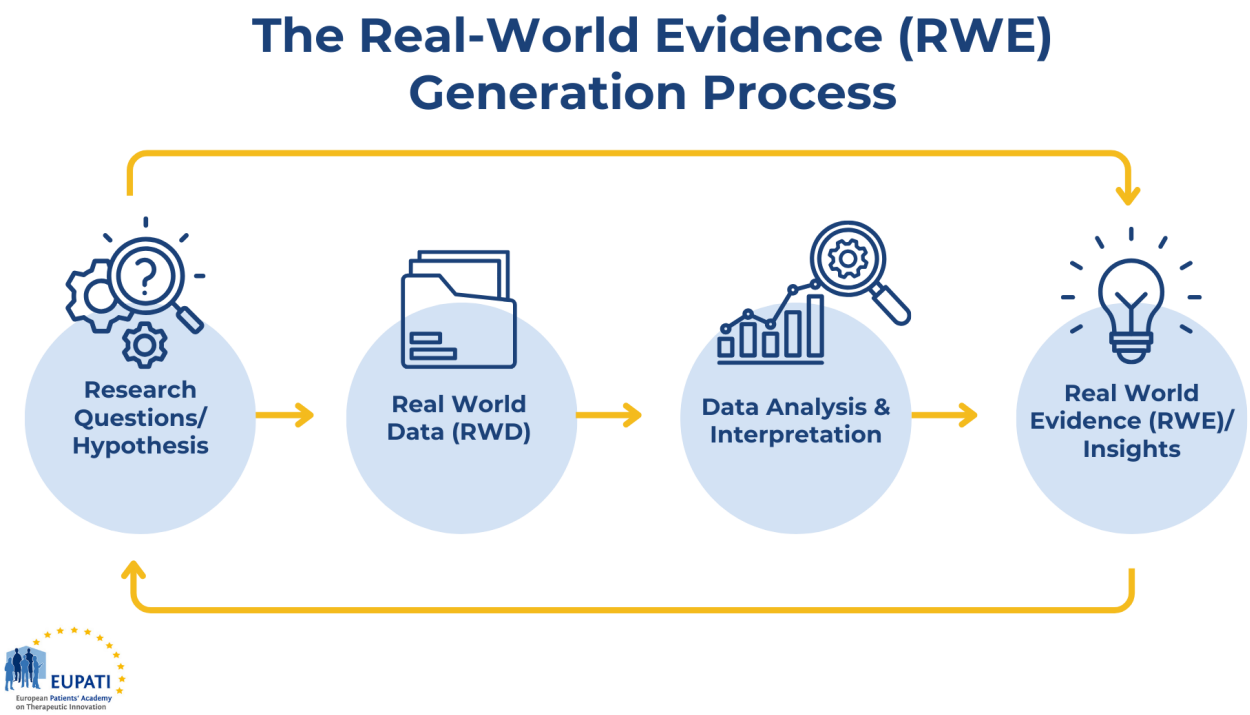

Figure 2: The Real-World Evidence Generation Process – Created by EUPATI

RWE can be generated from a large range of study designs and analytical methods (including quantitative and qualitative methods). The types of evidence generated from RWD are diverse and can range from characterising patients, conditions, or care delivery to developing economic models or estimating comparative effectiveness and cost-effectiveness. The choice of study design depends on the specific research question and its context. The following table highlights experimental, observational, and mixed study designs used to estimate the comparative effects of interventions based on RWD. Summaries of each design method are provided in Table 1 below[4] & [30].

Common Study Designs	
Cohort Study	A type of longitudinal observational study which collects data on a defined group of people over time. Cohort studies can be used to estimate the causal impact of exposures (including medical interventions) on clinical or other outcomes.

External Control Arm (ECA) for Clinical Trial	A special case of a cohort study in which the data for people with different exposures comes from different data sources. Typically, data on a new treatment from a clinical trial is compared to data from an external source, either a previous trial or real-world data, to form an indirect treatment comparison.
Pragmatic Randomised Controlled Trial	A pragmatic trial aims to measure the relative effectiveness of treatments in real-world clinical practice retaining the benefits of randomisation. Pragmatic randomised controlled trials may use real-world data in several ways including for recruitment or patient follow-up.
Other Observational Designs	
Quasi-Experimental Study	Aim to estimate causal effects of exposures using external variation in exposure across people or over time that is otherwise unrelated to the outcome. Examples of quasi-experimental methods include instrumental variable analysis, regression discontinuity, interrupted time series and difference-in-difference estimation.
Case-Control	A study that examines associations between outcomes and prior exposures by comparing people with an outcome of interest to those without the outcome. These generally do not allow for quantification of relative risk and are not used to estimate the relative effectiveness of interventions.
Cross-Sectional	Data are collected from a population or a representative subset of a population at one specific point in time (or over a short period) to examine associations between health status and exposures (use of interventions). The cross-sectional relationships revealed generally are not robust evidence of relative effectiveness, which requires longitudinal data.
Self-Controlled	Self-controlled, or 'within-subject', designs make use of variation in exposure status within individuals over time. These include case-crossover, self-controlled case series, and variants of these designs. They are most appropriate for transient exposures with acute-onset events (Hallas and Pottegard 2014).
Case Report /Case Series	A detailed report on a smaller number of patients, typically describing symptoms an unusual or new occurrence, including outcomes after a treatment.
Other Experimental Designs	
Population Enrichment RCT	Includes patients with characteristics typically under-represented or excluded from RCTs. Predictive modelling techniques may be applied to data generated from these studies to facilitate the estimation of relative effectiveness in a real-world population.
Cohort Multiple RCT (cmRCT) (also known as trials within cohorts)	A type of pragmatic RCT that uses a large cohort of patients as a source of participants for a variety of RCTs, providing a more generalisable study sample.
Comprehensive Cohort Study (CCS)	A type of pragmatic RCT that includes participants who do not consent to be randomised to the treatment group. This facilitates reduction in selection bias and improves the generalisability of study results.
Cluster RCT	Cluster RCTs randomise groups or clusters rather than individual participants as in traditional RCTs. This can reduce the possibility of contamination of the comparator (usual care) group.
Non-Randomised Controlled Trial	Any experimental study in which patients are allocated to different treatments using a method other than randomisation, such as clinician or patient preference.

Table 2: Study Designs Applicable to Use of RWD for

Comparative Effect Estimation [31]

[glossary_exclude]RWD vs. RCT Data: A False Dichotomy[/glossary_exclude]

Real-world data (RWD) and randomized controlled trial (RCT) data are often contrasted in literature, with RCTs frequently deemed superior due to their methodological rigor. RCTs eliminate confounding factors by including a control group and randomizing treatment allocation, thereby reducing bias from both known and unknown variables. In contrast, while RWD can yield accurate treatment effect estimates if confounders are properly adjusted, it is often challenging to identify and account for all relevant factors, increasing the risk of erroneous conclusions. Additionally, RCTs are designed prospectively, ensuring data quality and reliability. Conversely, RWD often derives from clinical data originally collected for patient care, which may lack the necessary details, particularly on drug toxicities and comorbidities, which regulators require to make informed benefit-risk assessments.

Despite these differences, the divide between RWD and RCT data is largely artificial. Certain RCTs, like pragmatic trials, integrate real-world conditions by implementing more inclusive eligibility criteria, minimizing study procedures, and leveraging existing healthcare infrastructure for outcome monitoring. Such trials are designed to emulate real-world settings, offering generalisable estimates of an intervention’s effectiveness. Similarly, studies using randomization within RWD, known as “randomized RWD” can be seen as enhanced forms of RWD, though they may still face issues like protocol deviations, which require adjustments

similar to those in typical RWD analyses.

There are instances where RCTs are neither feasible nor timely, and in such cases, RWE from observational RWD studies can be an invaluable alternative. However, caution should be exercised in relying solely on RWD for new therapies, given its vulnerability to bias. RCTs should remain the gold standard for assessing efficacy, but the focus should increasingly be on conducting trials capable of generating RWE. RWD studies can also serve as a benchmark for extending RCT findings. If an RWD analysis replicates RCT results, even when emulating the trial, it suggests that RWD may sufficiently approximate the effect estimates from the RCT. This allows for extending the RCT findings to broader patient populations or longer follow-up periods, an advantage given the impracticality of long-term RCTs across all patient demographics.

In summary, while RCTs and RWD have distinct strengths, the two can be complementary. Through rigorous RWD studies and RWE studies researchers and regulators can gain insights that are both scientifically robust and broadly applicable to populations [9]&[33].

[glossary_exclude]Advantages of RWE Studies Compared to Randomized Clinical Trials (RCTs)[/glossary_exclude]

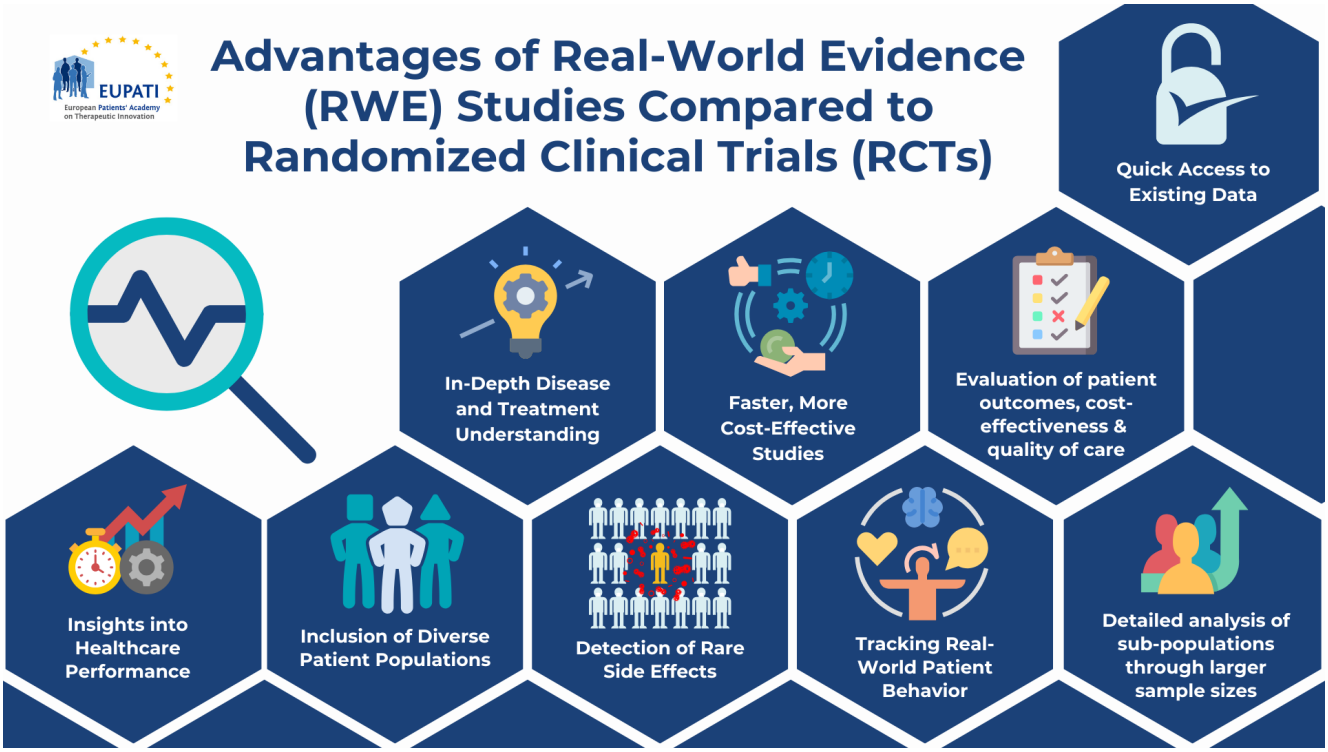

Figure 3: Advantages of RWE Studies compared to RCTs – Created by EUPATI

- **Time, Cost, and Resources**: RWE studies, depending on their design, generally require less time, fewer resources, and lower costs compared to RCTs. This is particularly advantageous in situations where faster research is needed, as they require less time for patient recruitment, enrollment, and study completion.
- **Rapid Data Access**: Data for RWE studies can be quickly accessed from existing databases and registries, allowing for more straightforward retrieval. This makes it easier to conduct large-scale studies without the lengthy data collection processes required by RCTs.
- **Understanding Disease and Treatment**: RWE is valuable for studying the natural history of diseases, such as their prevalence, incidence, unmet medical needs, treatment patterns, and standards of care. This allows for a deeper understanding of how diseases progress and how they are currently managed in Real-world settings.
- **Patient Outcomes and Health Economics**: RWE can be leveraged to evaluate patient outcomes and assess health

economics, providing insights into the cost-effectiveness and quality of care. This can help to make informed decisions about the value of treatments in clinical practice.

- **Insights into Healthcare Services**: RWE helps assess and understand current healthcare services, highlighting areas for improvement and providing a comprehensive view of Real-world healthcare performance.
- **Studies in Hard-to-Reach Populations**: RWE studies typically have fewer strict eligibility criteria compared to RCTs, meaning that a broader and more representative patient population can be included and research can be conducted in high-risk or vulnerable populations who are often excluded from traditional clinical trials (e.g., pregnant women, children, the elderly, or those with multiple comorbidities). This helps fill critical gaps in knowledge regarding treatment responses in these groups.
- **Rare Side Effects**: RWE studies are particularly useful for identifying less common side effects, as RCTs often have smaller sample sizes and shorter durations, which makes it more difficult to observe infrequent adverse events. The larger and more diverse sample sizes in RWE studies improve the detection of these rare side effects.
- **Tracking Real-World Patient Behavior**: RWE studies can provide valuable insights into Real-World patient behaviors, such as adherence to treatment, lifestyle factors, and other variables that may influence health outcomes. These aspects are often difficult to capture in RCTs, but RWE allows for a more complete understanding of how patients interact with healthcare treatments in daily life.
- **Large Sample Size for Sub-population Analyses**: The larger sample sizes typically seen in RWE studies facilitate the ability to analyse sub-populations, allowing for more detailed investigations into how

treatments affect different groups. This is especially useful for analyzing treatment effects in smaller or less common patient groups, providing insights that may be missed in smaller RCTs. Additionally, large sample sizes contribute to greater generalizability of the findings to broader, real-world patient populations[9],[33]&[34].

[glossary_exclude]Real-World Evidence in Action: How COVID-19 Vaccine Effectiveness Studies Shaped Global Policy and Public Health Strategy[/glossary_exclude]

During the COVID-19 pandemic, RWE played a critical role in assessing the real-world effectiveness of COVID-19 vaccines. Data from various sources, was analysed to track vaccine efficacy, monitor side effects, and evaluate outcomes in different patient populations. These studies informed policy decisions, including booster dose recommendations, specific guidance for at risk populations, and the global COVID-19 response.

One notable study, titled " *Real-Word Effectiveness of Global COVID-19 Vaccines Against SARS-CoV-2 Variants: A Systematic Review and Meta-Analysis*", published in *Frontiers in Medicine* in 2022, examined the effectiveness of vaccines using data from global cohorts and clinical trials. This study exemplifies how RWE was crucial in shaping clinical and regulatory decisions. Below is a summary of the study's key findings and its role in guiding policy decisions related to vaccine efficacy, booster doses, and public health strategies.

Summary of the Study:

- **Objective**: The study aimed to assess the real-world effectiveness of COVID-19 vaccines in preventing symptomatic infection, severe disease, and death, drawing on evidence from a variety of global cohorts.
- **Data Sources**: The study used a wide array of data sources, including EHRs, insurance claims, clinical trial data, and registry data from countries around the world.
- **Key Findings**:
  - COVID-19 vaccines were found to be highly effective in preventing symptomatic infection, severe disease, and death across different population groups.
  - Effectiveness varied depending on factors such as age, underlying health conditions, and the presence of SARS-CoV-2 variants like Delta and Omicron.
  - The vaccines also showed a decrease in effectiveness over time, highlighting the importance of booster doses for maintaining high protection levels, especially against emerging variants.
  - Evidence demonstrated that vaccines were effective in both general populations and vulnerable groups, such as the elderly and those with underlying health conditions.

[glossary_exclude]Role of RWE in Assessing Vaccine Effectiveness:[/glossary_exclude]

- **Data Integration**: RWE played a crucial role in providing a comprehensive picture of how the vaccines performed in diverse, real-world settings. Unlike controlled clinical

trials, RWE integrated data from large-scale populations, including various age groups, pre-existing health conditions, and different geographic regions.

- **Monitoring Variants**: The study helped track the vaccines' effectiveness against new variants of concern (VOCs), such as Delta and Omicron. This is especially important for adjusting vaccination strategies to respond to shifting viral mutations.
- **Impact of Booster Doses**: The study's findings, showing a decline in effectiveness over time, played a pivotal role in the recommendation for booster doses. This data directly informed policies in multiple countries, ensuring that vulnerable populations received additional protection.
- **Real-World Guidance**: Through RWE, the study also provided insights into how the vaccines performed in groups that were often underrepresented in clinical trials, such as elderly individuals and those with underlying health conditions. These findings were crucial for informing public health recommendations tailored to at-risk populations.

How the Study Informed Policy Decisions:

- **Booster Dose Recommendations**: The Real-world evidence of reduced vaccine efficacy over time supported the widespread recommendation for booster doses. Health authorities, including the World Health Organization (WHO) and national health agencies, used this data to advocate for additional vaccine doses to sustain high protection levels against symptomatic infection and severe outcomes.
- **Guidance for At-Risk Populations**: The study highlighted the need for more robust vaccination strategies targeting vulnerable populations, such as the elderly, immunocompromised individuals, and those with chronic conditions. The evidence guided specific recommendations

for prioritizing these groups for early and additional doses.

- **Global COVID-19 Response**: The findings from this study, based on Real-world data, were instrumental in shaping the global COVID-19 response. Governments and health organisations used this information to adjust public health strategies, allocate vaccines more efficiently, and communicate the evolving need for boosters in response to new variants.

Conclusion:

This study underscores the critical role of Real-world evidence (RWE) in assessing the effectiveness of COVID-19 vaccines and providing actionable insights for policy decisions. By integrating data from diverse populations and settings, RWE helped refine vaccine strategies, shape booster dose policies, and ensure that vulnerable populations received timely protection. The study's findings were pivotal in guiding the global response to the pandemic and optimizing vaccination efforts worldwide[31].

[glossary_exclude]Ensuring Transparency and Rigor in Real-World Evidence (RWE) Studies: Best Practices and Regulatory Initiatives[/glossary_exclude]

Maintaining transparency throughout each stage of a study is essential for ensuring reliability and fostering trust in real-world evidence (RWE).



Figure 4: A Step-by-Step Guide to Ensuring Rigor and Transparency in Real-World Evidence Studies – Created by EUPATI

Key steps include careful planning, responsible data handling, clear documentation, and accessible reporting:

Study Planning

- **Defining the Research Question**:
  - Clearly outline the research question, specifying:
    - Key study variables: population, interventions/exposures, outcomes, covariates
    - Subgroups, with validation of subgroup categories where relevant
    - Target quantities (e.g., disease prevalence, effect on survival)
  - Choose patient-centered outcomes that reflect well-being function, or lifespan, using validated outcome sets (e.g., COMET database).
  - For non-randomized studies, provide a rationale

for the approach due to the absence of randomized evidence or trial limitations.

- **Pre-Specifying Study Conduct**:
    - Protocols should include objectives, data sources, design, and analytical methods for planned analyses, aiming for transparency and reproducibility (e.g. the HARPER tool, the START-RWE tool, or the Principled tool[,[28], [29] & [31]].
    - Pre-specifying analysis plans mitigate bias risk by discouraging selective analysis; ideally, publish protocols on accessible platforms (e.g., ClinicalTrials.gov, ISRCTN, EU-PAS, OSF) [22],[23],[28] &[29]. Further guidance on registration of study protocols is provided by the NICE's Advice service also provides advice on how technology developers can make the best use of real-world data as part of their evidence-generation plans.
- **Addressing Equality and Diversity**:
    - Ensure study design considers relevant equality or diversity issues.
- **Data Source Selection**:
    - Justify the choice of data sources based on their suitability and provenance, using a systematic, transparent search. Strategies include:
        - Defined dataset criteria and expert consultation
        - Flow diagrams showing data source inclusion/exclusion
    - Comply with data protection laws, securing necessary approvals and permissions.

A positive example is the Health Data Research UK Innovation Gateway, a platform that provides researchers with access to a

wide range of health data from sources like the NHS and academic institutions. It aims to accelerate medical research and innovation by facilitating responsible, data-driven studies while ensuring privacy and ethical standards.

- **Primary Data Collection**:
    - When needed, collect primary data through observational studies, surveys, or quality-of-life questionnaires, and consider sampling methods (e.g., random, stratified, purposive) to minimize participant burden.

Study Conduct

- **Design and Analytical Methods**:
    - Choose designs and methods appropriate to data characteristics, considering outcome distribution, sample size, data structure, population heterogeneity, and data type (cross-sectional or longitudinal).
    - Conduct diagnostic checks to assess study robustness through sensitivity analyses.
- **Minimizing Risk of Bias**:
    - Address potential threats to validity, such as selection, information, and confounding bias, particularly in comparative effect studies.
- **Quality Assurance**:
    - Use validated code and perform quality checks as recommended in the UK Government’s Aqua Book and the Office for National statistic’ s Quality Assurance of Code for Analysis and Research guidelines.

Reporting

- **Comprehensive Documentation**:
  - Sufficient documentation should support study reproducibility, aligning with STROBE, RECORD, RECORD-PE, ESMO-GROW checklists as applicable, and making operational definitions and methodologies clear.
- **Data Curation and Analysis**:
  - Report data curation and analysis steps, including software and code details. Make code publicly available (e.g., GitHub) where feasible, while protecting privacy and intellectual property.
- **Study Methods**:
  - Define study design variables, follow-up periods, and confounder identification processes. Use study diagrams or templates to illustrate timelines.

The Reproducible Evidence: Practices to Enhance and Achieve Transparency (REPEAT) initiative’s project page serves as a positive example, offering access to both the research paper and design diagram templates.

- **Statistical Reporting**:
  - Provide central estimates, precision measures, and results for main and sensitivity analyses, noting which analyses were pre-specified. Report unadjusted results where confounding adjustments are applied.
  - Ensure that information in figures and tables cannot inadvertently identify patients. The Office for National Statistics has guidance on maintaining confidentiality when disseminating health statistics.

Interpreting and Communicating Results

- **Result Interpretation**:
  - Discuss data, design, and analysis limitations, with clear implications for the findings.
- **Clear Communication**:
  - Present study information accessibly, explaining terms and avoiding jargon. Clearly label tables and graphs for easy interpretation[28] & [29].

RWE in Regulatory Approval

- The FDA’s Oncology Center of Excellence Real World Evidence (OCE RWE) Program launched guidance in June 2023 as part of its Quality, Characterization, and Assessment of Real-World Data (QCARD) Initiative. This guidance outlines key design and data source elements for oncology RWE studies to enhance the quality of initial study proposals.

- The European Medicines Agency (EMA) has offered guidance on data source selection through its RWE Methodology Working Party.

- Another EMA initiative aimed at improving RWE quality is the Data Analysis and Real World Interrogation Network (DARWIN EU). This pan-European, federated data network will facilitate the exchange of real-world data for healthcare delivery, policymaking, and research across Europe.

[glossary_exclude]RWE in HTA[/glossary_exclude]

- CADTH, in collaboration with Health Canada, has developed guidance for Reporting Real-World Evidence (RWE)[24], which lays the foundation for RWE's integration into regulatory approval and HTA in Canada. The guidance emphasizes transparency in reporting RWE studies, ensuring regulators and HTA bodies have the necessary information to evaluate their suitability for decision-making.

- The European Network for Health Technology Assessment (EUNetHTA) has introduced the Registry Evaluation and Quality Standards Tool (REQueST), which assists registry owners in maximizing data quality and ensuring that registry-based RWE can be utilized for HTA and regulatory purposes.

- The **Structured Process to Identify Fit-For-Purpose Data (SPIFD)** and its updated version, SPIFD2, provide data feasibility assessments with step-by-step decision tools and templates to aid in the identification, selection, and rationalization of fit-for-purpose RWD sources.

- **NICE's Data Suitability Assessment Tool (DataSAT)**: The introduction of the DataSAT by NICE emphasizes the growing expectation for transparency in the selection of data sources for RWE studies used in HTA. This tool forms part of NICE's broader RWE framework, ensuring

that the data supporting HTA decisions is robust and fit for purpose[17], [18] & [25].

[glossary_exclude]Key Considerations for Addressing Bias in RWD[/glossary_exclude]

RWD's credibility faces challenges due to biases from lack of randomization, issues in data quality, and risks of misleading findings from data mining. These concerns have slowed progress in harnessing RWD despite longstanding data capabilities in healthcare[21].

Despite these challenges, data mining, when properly applied, unlocks significant potential in RWD. It enables key advancements in healthcare, such as understanding treatment patterns, predicting outcomes, identifying high-risk patients, ensuring drug safety, and evaluating cost effectiveness:

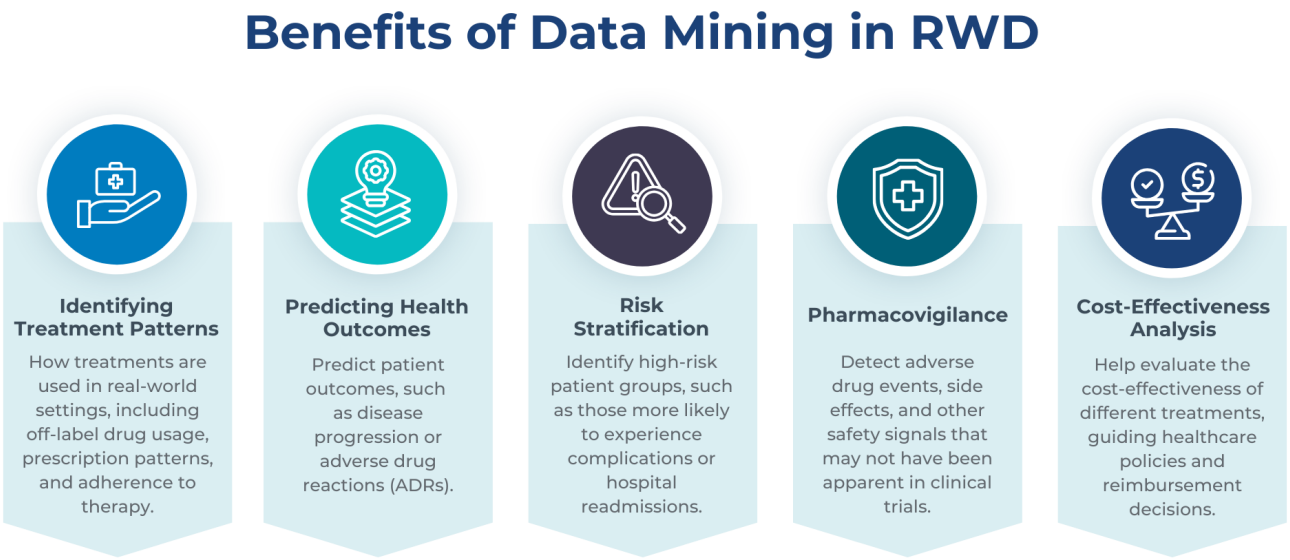

Figure 5: Benefits of Data Mining in Real-World Data – Created by EUPATI

To achieve these outcomes, data mining in RWD applies a

variety of advanced techniques:

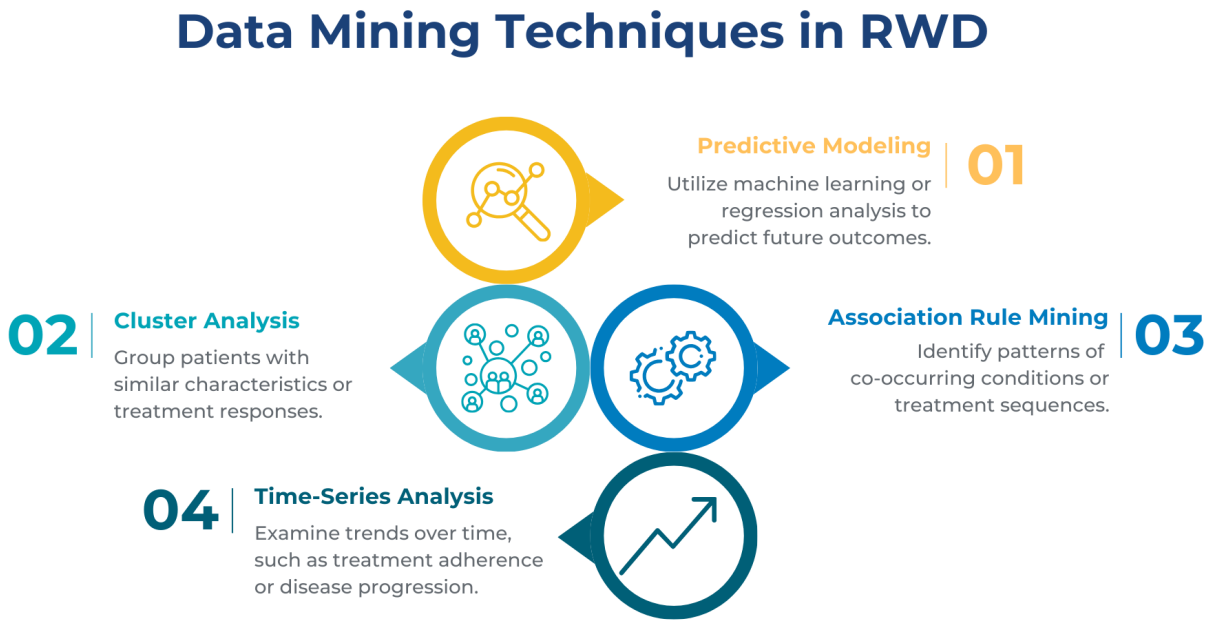

Figure 6: Data Mining Techniques in Real World Data – Created by EUPATI

RWD generation also faces specific challenges in certain contexts, such as rare diseases, medical devices, interventional studies, and digital health technologies.

Key challenges include:

- **Rare Diseases**[43]**:**
  - Challenges in systematic identification of target populations.
  - Small sample sizes, often requiring integration of data from various sources with differing data models and collection methods.
  - Lack of standardized common data elements.
  - Significant variation in the natural progression of disease.

- **Medical Devices and Interventional Studies**[44]**:**
  - Limited national datasets that integrate device use and patient outcomes.

- Insufficient detail in routine data to track specific devices (e.g., unique device identifiers) or procedures.
- Difficulty identifying appropriate comparators due to technology evolution and learning effects over time.

- **Digital Health Technologies**[45]**:**
  - Challenges in evaluating rapidly evolving technologies and their integration into healthcare systems.
  - Lack of consistent frameworks for assessing their effectiveness and safety.
  - Variability in how digital health data is collected, stored, and analysed.
  - Ethical concerns regarding data privacy and patient consent.

Improvements in data systems, including registries and electronic health records, can help to maximise reliability and minimise bias[42].

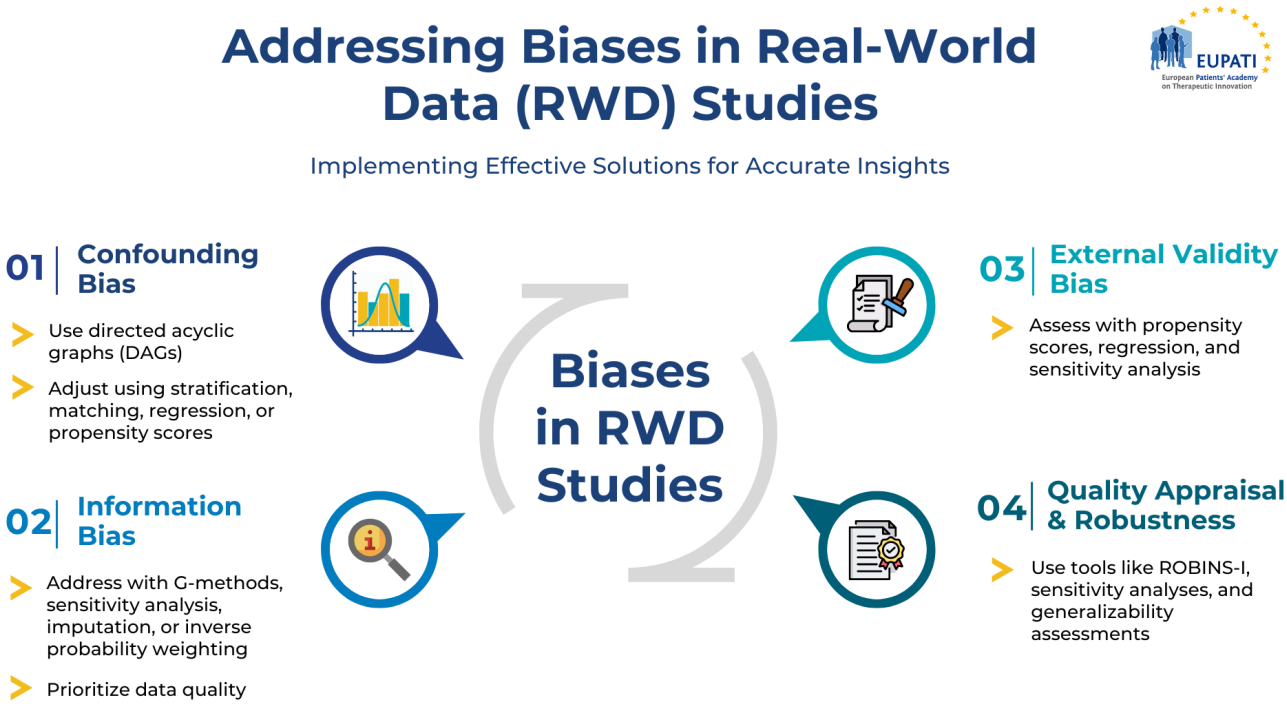

Figure 7: Identifying and Managing Bias in RWD Studies – Created by EUPATI

Addressing Confounding Bias

To minimise confounding bias, researchers should systematically identify and select confounders based on literature and expert input, focusing on time-varying confounders when relevant. Establishing causal relationships through directed acyclic graphs is recommended, and inappropriate adjustment (e.g., over-adjusting for mediators or colliders) should be avoided. Machine learning techniques may help identify covariates in large datasets, provided assumptions are well-justified.

Adjusted comparisons that follow clear causal assumptions are preferred over unadjusted ones. Approaches to adjust for observed confounders include stratification, matching, multivariable regression, and propensity scores. Propensity score methods, such as matching or stratification, facilitate transparency, while advanced methods like marginal structural models are better suited for time-varying confounders.

Managing Information Bias

Data quality issues such as missing data, measurement error, and censoring can introduce bias. Informative censoring, where dropout correlates with treatment or outcomes, can be addressed with G-methods or sensitivity analysis. For missing data, methods like imputation or inverse-probability-of-treatment weighting are useful, while sensitivity analysis can assess the impact of missing data that do not follow random patterns. High specificity should be prioritized to reduce measurement error and misclassification biases.

Ensuring External Validity

External validity bias, which arises from differences between study samples and target populations, should be assessed by

comparing patient characteristics and contexts. Statistical tools, such as propensity scores or regression methods can adjust for these differences, though success depends on adequate overlap between the sample and target populations. Sensitivity analyses are essential to explore potential biases when transferring findings to new settings populations [9] & [28].

Quality Appraisal and Robustness

At the planning stage, researchers should identify potential biases and outline methods to address them, using tools like ROBINS-I and GRACE for systematic quality appraisal[25], [26], [27] & [28]. Although ROBINS-I is a preferred tool for assessing non-randomized studies, it does not cover all bias types, and statistical uncertainty in non-randomized studies often understates true risk. Developers should assess how well study results generalize to their respective healthcare settings (e.g., NHS in the UK, or other relevant systems), by examining differences in patient populations or care settings, backed by a data suitability assessment.

To ensure robustness, sensitivity analyses should be conducted, comparing adjusted and unadjusted analyses to clarify any deviations and confirm that results remain consistent across different assumptions. This approach strengthens confidence in the findings by demonstrating resilience to potential biases and assumptions.

In summary, the rigorous design and assessment of RWD studies, addressing both internal and external biases, are essential for producing valid, generalizable insights[28].

Patient Involvement in Real-World Evidence Assessment

Each type of data source has some general strengths and weaknesses, but the value of a given research question will depend on the characteristics of the specific data.

Patients can contribute to the discussion on RWE relevance and quality as they can confirm whether the data under discussion is relevant from a patient perspective.

Patients who are involved in research discussions should also ask for a specified analysis plan to ensure that patient-relevant aspects are not distorted or otherwise interpreted in a biased way. This requires patients to have good knowledge about:

- The purpose of the primary data collection and how it aligns with patient-relevant aspects.
- Whether an available dataset is sufficient to fully answer the scientific question(s).
- Analysis methods that will be used.
- Limitations of what can be concluded from the dataset.
- Risk of cherry-picking existing datasets and analysis with already-known results.

The **European Medicines Agency (EMA)** prioritizes patient involvement through various initiatives that span the entire regulatory lifecycle. EMA's commitment includes active collaboration with patient organizations to inform regulatory decisions, integrating patient-reported outcomes **(PROs)** into the assessment of treatment efficacy, and ensuring patient representation on scientific committees, such as the **Committee for Medicinal Products for Human Use (CHMP)**. These efforts aim to reflect the perspectives and priorities of patients in both the pre-market and post-market phases of drug development. For further information on how EMA engages with patients and the

public, visit EMA Patient Involvement[49].

Similarly, the **FDA's Center for Drug Evaluation and Research (CDER)** supports patient involvement through its Patient-Focused Drug Development (PFDD) program. This initiative aims to systematically incorporate patient experiences, needs, and perspectives into the drug development and evaluation process. The program helps to ensure that patient insights are considered in regulatory decision-making by capturing the lived experiences of patients with their conditions. Key aspects of the PFDD program include gathering patient input through structured methods, improving enrolment and reducing participant burden in clinical trials, and developing methods to assess patient preferences regarding treatment benefits and risks[6].

For basic information on types of evidence and how they are used in evaluations, please see the EUPATI Course ''*Concepts of Evidence-Based Medicines and Outcomes Research* (https://learning.eupati.eu/course/view.php?id=7)''.

How can patients play a critical role in shaping Real – World Data (RWD) and Real – World Evidence (RWE)?

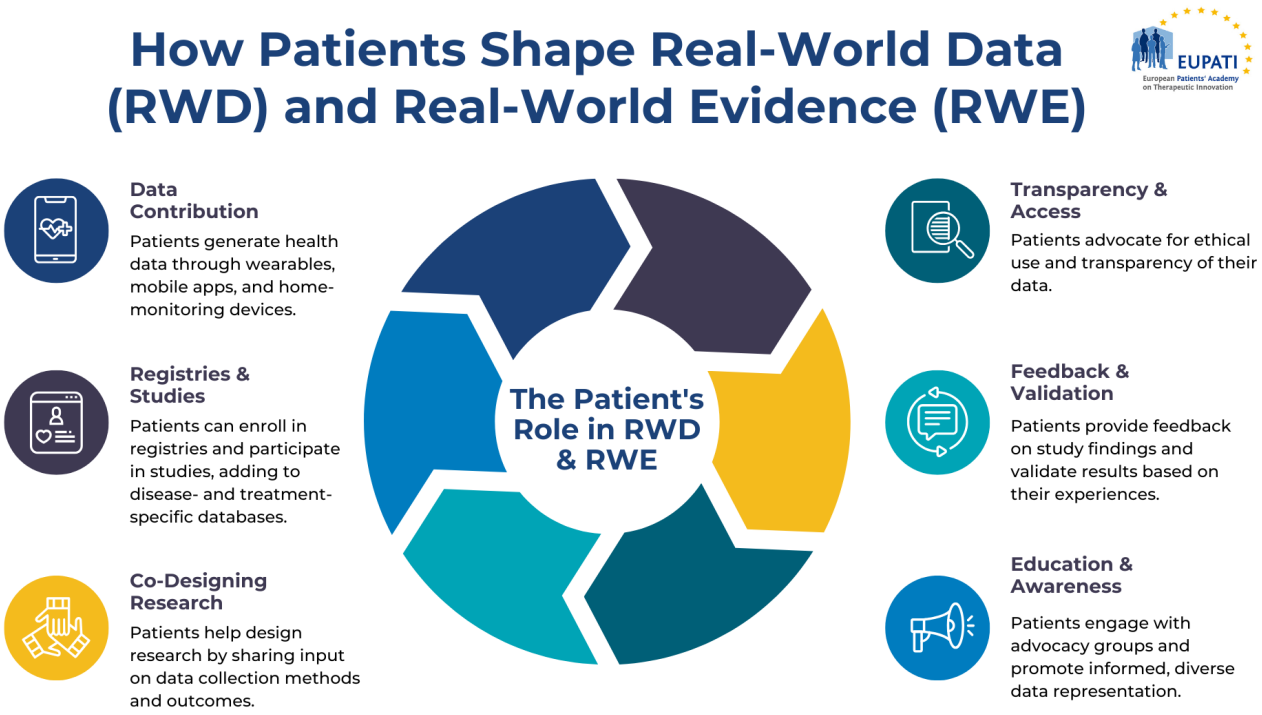

Figure 5: The Patient's Role in RWD & RWE – Created by EUPATI

1. Data Contribution:

– **Patient-Generated Health Data (PGHD)**:

Patients can play an essential role in managing their health by utilizing tools such as electronic health records (EHRs), patient portals, wearable technology, and home-monitoring systems. The shift from paper-based to digital health records has significantly enhanced the efficiency of data storage, retrieval, and sharing, while simultaneously empowering patients to become more engaged in their healthcare. Furthermore, the increasing availability of standardized digital tools and platforms has simplified the process for patients to track important health metrics, including symptoms and medication usage. At least some of these tools can also be used in randomized controlled trials (RCTs) to enhance data collection.

PGHD are increasingly being generated through the following methods:

- Integrated Health Platforms: Smartphone ecosystems now often feature health platforms that aggregate data from various fitness and health applications, creating a comprehensive view of the user's health.

- Wearable Devices: Devices continuously collect real-time data on physiological parameters like heart rate, ECG, cuff-less blood pressure, Oxygen Saturation (SpO_2), blood sugar levels, sleep patterns, and physical activity.

- Smart Clothing: Garments equipped with sensors that also monitor physiological indicators like heart rate, respiration, and muscle activity.

PGHD can also be generated using the following home-monitoring systems:

- Blood Pressure Monitors: Useful for individuals managing hypertension or cardiovascular conditions, allowing for regular monitoring at home.

- Glucose Meters: Essential for diabetics to routinely check their blood sugar levels.

- Pulse Oximeters: Important for those with respiratory conditions to measure blood oxygen saturation.

- Electronic Scales: Often used by patients with heart failure or those undergoing treatments where weight monitoring is crucial.

- Smart Thermometers: Digital thermometers that monitor and record body temperature, often linking to apps for easier tracking.

Through the use of these advanced technologies, patients are better equipped to monitor and manage their health, contributing to more informed and engaged healthcare decisions.

– Patient-Reported Outcomes (PROs):

Patients can offer direct feedback on their health status, symptoms, and quality of life through surveys and digital tools, adding depth to RWD by capturing insights that may not be evident in clinical environments. Beyond biometric data, PROs can also include lifestyle habits, daily activities, and other health-related data, offering a more holistic view of patient's health.

This feedback can be collected through various methods:

- Health Diaries: Patients can maintain regular logs of their symptoms, diet, medication use, and other health-related behaviors, providing detailed, ongoing insights into their health status.

- Surveys & Questionnaires: Tools like the Patient-Reported Outcomes Measurement Information System (PROMIS) offer standardized measures to assess physical, emotional, and social well-being from the patient's perspective.

- Mobile Apps: Many wellness apps allow users to record data on aspects such as mood, diet, menstrual cycles,

and pain levels. These apps, often integrated with wearable devices, also use gamification elements to motivate patients to reach health goals, participate in challenges, or track progress, which encourages consistent data logging. Additionally, social features like sharing achievements and joining group challenges can further boost patient engagement and data collection.

PROs are increasingly valued for their potential to enhance care and research, particularly in areas like chronic disease management, mental health, and preventive care[7].

To deepen your understanding of Patient-Reported Outcomes (PROs), explore the resources available in the EUPATI Open Classroom and EUPATI Toolbox.

2. Participating in Registries and Studies:

– **Patient Registries**: Patients can enroll in disease-specific or product- (e.g. treatment or device) specific registries, where their health data (such as clinical information or biological samples stored in bio-banks) is systematically collected and analyzed to generate RWE on long-term outcomes. These registries help evaluate treatment effectiveness, monitor disease progression, ensure product safety, and identify high-risk patient populations and unmet medical needs.

Registries can either be hospital-based, collecting data from patients diagnosed with and treated for a specific disease at a single hospital or across multiple hospitals, or population-based, gathering information from all individuals within a defined geographic region. For instance, the European Cystic Fibrosis Society (ECFS) Registry compiles demographic and

clinical data from cystic fibrosis (CF) patients across Europe. This data is used to monitor and analyze various aspects of CF and its treatment in EU countries, enhance care standards, support CF-related epidemiological research, and accelerate the development of public health policies.

A major advantage of registries is their adaptability, allowing for extensive data collection that can be used for multiple studies. Over time, these registries can be modified to include new variables or outcomes of interest. Additionally, they have been used to better understand how physicians make clinical decisions[8]& [9].

– **Observational Studies**: Patients can participate in observational studies or post-marketing surveillance programs, helping to generate RWE about the real-world performance of treatments and interventions.

Observational studies can take several forms, including:

- **Cohort studies**, which evaluate disease incidence, aetiology, risk factors, natural disease progression, prognosis, and treatment outcomes.

- **Cross-sectional studies**, which examine disease prevalence and outcomes by assessing a single group of patients at a specific point in time, with treatments and outcomes studied concurrently.

- **Case-control studies**, which focus on a single outcome and explore potential causes, making them particularly useful for investigating rare conditions or diseases with long latency periods between exposure and onset.

Data for RWD studies can be gathered either **prospectively**, where new data is collected from RWD sources,

or **retrospectively**, where existing data from RWD sources is analysed. A key feature of RWE studies is that treatment is administered in accordance with marketing authorisation, physician discretion, and national or regional treatment guidelines, rather than following a pre-specified protocol, as seen in randomized controlled trials (RCTs).

In many cases, prospective, multicenter observational studies are conducted as part of routine clinical care, and these are referred to as **pragmatic clinical trials**. Pragmatic trials are designed to provide evidence that supports clinical or policy decisions by assessing the effectiveness of drugs or medical products in real-world practice settings[9]. While RCTs remain the primary standard for marketing authorization applications, pragmatic clinical trials (PCTs) are increasingly recognized for their role in providing valuable real-world evidence. Regulatory agencies like the FDA and EMA are open to using PCTs as supplementary evidence or as part of an adaptive approval process[48].

3. Co-Designing Research and Data Collection:

– **Patient Engagement**: Patients should be actively involved at every stage of designing RWD studies. Working alongside researchers, healthcare providers, and policymakers, patients play a crucial role in defining the research framework, selecting suitable data collection tools (such as patient-reported outcome instruments), identifying meaningful outcomes, and advocating for broader participation[10].

The international Core Outcome Measures in Effectiveness Trials (COMET)[46] Initiative, aiming to enhance the development and implementation of standardized endpoints in clinical research, referred to as “Core Outcome Sets” (COS), is a

bright example of effective patient engagement. COS defines the essential outcomes that must be measured and reported in clinical trials for a particular condition. Involving patients in developing COS ensures research is aligned with patients' priorities. The COMET PoPPIE Working Group supports global COS researchers in integrating patient perspectives. Patients can also contribute to COS studies through activities like surveys or consensus meetings.

– **Advisory Roles**: Patients can serve on advisory boards or committees that guide the development of RWE studies, ensuring that the research addresses patient priorities and concerns.

PCORnet is another positive example of the long-sought ideal of a comprehensive research ecosystem—a fully integrated network that seamlessly combines extensive, highly representative health data with research expertise and patient insights, ensuring they are accessible from the outset. PCORnet ensures patient and caregiver voices are central to research, with active roles on the Steering Committee and leadership throughout the process[11].

4. Advocating for Data Transparency and Access:

– **Data Governance**: Patients can advocate for transparent data practices, ensuring that their data is used ethically and that they have control over how their information is shared and analysed. Data governance arrangements must balance the interests of various stakeholders—advancing medical knowledge through evaluation and research while safeguarding individual privacy.

– **Informed Consent**: Patients can demand clear communication about how their data will be used in RWE studies, including

potential risks and benefits of participation.

– Privacy and Regulatory Challenges: Privacy concerns are governed by regulations such as the General Data Protection Regulation (GDPR) in the European Union (EU) and the Health Insurance Portability and Accountability Act (HIPAA) in the U.S. In the EU, the legal and ethical framework for health research is both complex and inconsistent, creating significant obstacles to cross-border collaboration. This complexity can particularly disrupt clinical trials, complicate the secondary use of data, and reduce the level of protection for research participants, referred to as "data subjects" under the GDPR[13]&[14].

Given the fragmented implementation and interpretation of the GDPR at the national level, the European Commission is taking steps to address this gap through the **'Towards European Health Data Space (TEHDAS2)'** project, which engages EU citizens in discussions about the use of health data. As a result, twelve key recommendations were developed to guide citizen involvement in the European Health Data Space (see Table 5 below for a summary). These recommendations are organised around three core concepts:

1) The Data Relationship,

2) The Power Balance, and

3) A Citizen-Powered Framework.

The Data Relationship
1. Citizens should be able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged.
2. Citizens should have access to their data and know how they are used for secondary purposes.
3. Citizens' values should inform what is beneficial to individuals and what constitutes the common good.
4. Decision-making processes should rely on a plurality of views and actors to increase their trustworthiness.
5. Citizens should be given the opportunity to be involved in the lifecycle of health data.
The Power Balance
6. Citizens should be provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value their ability to exercise control.
7. The protection of individuals' identity should be ensured; citizens perceive this as one of the best ways to balance the harms and benefits of the secondary use of health data.
8. Data users' intentions should be transparent and in line with purposes of citizens' support.
9. Accountability could be enhanced through transparent and stronger mechanisms.
10. Good IT solutions should be fostered to protect citizens' data, beyond having a strong legal framework in place.
A Citizen-Powered Framework
11. Stakeholders should respect principles that align with citizens' ethical values.
12. There should be a framework that facilitates the secondary use of health data for purposes and benefits that citizens support, while minimizing the potential risks that citizens' identify.

Table 2: Recommendations on how to engage citizens in the European Health Data Space[13]

Several organisations, including professional societies, regulatory agencies, and research consortiums, are simultaneously developing guidelines in the form of checklists

to ensure that data is fit for regulatory decision-making. A key advancement in promoting transparency in RWE strategy came with the release of guidance from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the International Society for Pharmacoepidemiology (ISPE) Special Task Force. These recommendations focus on the data collection process, data quality and reliability, and follow the fit-for-purpose principle. Among the key suggestions for establishing international standards to assess the acceptability of RWD governance practices are:

- **Data Access Agreements**: Establishing clear and transparent criteria for data access, which should be approved by ethical and institutional review boards.

- **Pre-registration of RWE Studies:** It is advisable to pre-register RWE studies, detailing both the study protocol and the plans for data collection and analysis in advance.

- **Minimum-Quality Criteria**: Minimum-quality criteria for the RWD used in RWE studies should not be disregarded at the time of publication.

- **Quality Thresholds**: Implementing minimum-quality thresholds, or several graded quality thresholds, for data governance checklists is recommended. However, flexibility may be necessary when dealing with limited RWD[12], [17], [20].

5. ## Feedback and Validation:

– **Data Validation**: Patients can help validate the findings

from RWE studies by providing feedback on whether the results align with their experiences and outcomes in real-world settings.

– **Feedback Loops**: Continuous patient feedback can improve the relevance and accuracy of RWD, leading to more robust and patient-centred RWE[15]&[16]. Feedback loops between patients and researchers can refine RWE studies, allowing for adjustments based on real-world applicability and aligning study findings with patient experiences.

Patient feedback loops can refine RWE studies across multiple stages:

1. **Research Question**: Patient input, gathered through qualitative data or existing studies, can identify information gaps and ensure the study focuses on relevant questions that address unmet needs.
2. **Protocol Development**: Collaboration with patients to develop conceptual frameworks and patient-journey maps can help refine study design, including subpopulation characteristics and outcome selection.
3. **Data Sources**: Patient registries and experience mapping can highlight critical data gaps and improve data quality.
4. **Findings Translation**: Patient feedback can ensure findings are disseminated in a way that aligns with how patients access information, addressing study limitations and guiding future research agendas.

– **Collaboration for Data Integrity**: Data stewards (such as healthcare systems, vendors, and governments) play a key role in ensuring the accuracy, completeness, and reliability of real-world data (RWD). By facilitating collaboration between stakeholders, including patient representatives, clinicians and researchers, they enable the identification of missing variables and confounders that can affect study outcomes. This type of collaboration could help create a more

holistic understanding of patient needs and improve the applicability of RWE. Engaging diverse healthcare professionals, including physicians, nurses, and medical coders, could also ensure that study designs are grounded in comprehensive, patient-centred insights[16].

[glossary_exclude]For more detailed information on clinical trials, protocol design, and patient involvement, explore the following resources:

- The Clinical Trials and Trial Management course of the Clinical Development Module, in the **EUPATI Open Classroom**,
- The Mini-course Starter Kit – Protocol design, in the **EUPATI Toolbox**,
- The Guidance for patient involvement in ethical review of clinical trials, in the **EUPATI Toolbox,**
- The following dedicated **EUPATI Toolbox articles**, that can provide useful, real-world examples of patient involvement in clinical research:
- Patients Involved – Clinical trial design,
- Patients Involved – Patient input into breast cancer study design,
- Patients Involved – Patient feedback on a paediatric CML study,
- Patients Involved – Direct patient insight on Lupus,
- Patients Involved – DevelopAKUre for AKU,
- Patients Involved – An ultra orphan disease – EUPATI Toolbox,
- Patients Involved – Collaboration in an oncology trial,
- Patients Involved – HIV R&D Collaboration,
- Patients Involved – HIV patients actively involved in multiple phases,
- Patients Involved – HIV DUET Phase III trials,
- Patients Involved – Patient organisations' input on a rare disease registry,
- Patients Involved – Patient organisation promotes Phase

I clinical trial – EUPATI Toolbox,
- Patients Involved – Using user research – EUPATI Toolbox,
- Patients Involved – Patient friendly informed consent – EUPATI Toolbox,
- Patients Involved – Informed Consent Feedback,
- Patients Involved – Patient feedback on a plain language summary of results
- Patients Involved – Informed consent form: Writer's guide[/glossary_exclude]

6. Education and Awareness:

– **Patient Education**: By being informed about RWD and RWE, patients can better understand the importance of their data and actively engage in discussions about healthcare research and decision-making.

– **Community Involvement**: Patients can engage with patient advocacy groups and community organisations to promote the collection and use of RWD/RWE that reflects diverse patient populations and reduces disparities in RWE data representation. [19].

In summary, patients are essential to the development of RWD and RWE as they actively contribute their own health data, engage in research, and participate in shaping study designs. Their input helps ensure that healthcare research reflects real-world experiences and needs, promoting more relevant and effective health interventions.

Furthermore, by advocating for transparency and ethical standards in data use, patients foster a collaborative environment that strengthens the overall healthcare system, ultimately leading to improved health outcomes, policies,

treatment guidelines, and innovations, that better serve diverse populations.

The Path Forward

Artificial intelligence (AI) techniques, particularly machine learning (ML), have been critical for analysing unstructured data from sources like electronic health records (EHRs), social media, mHealth apps, and wearables. During the COVID-19 pandemic, ZOE's mHealth app, built on insights from the PREDICT study, leveraged ML to analyse RWD from millions of at-home participants, providing rapid, near-real-time insights on COVID-19 symptoms, trends, and vaccine efficacy when direct clinical data was limited. This approach sets a scalable model for remote health research[41]. Additionally, ML has proven effective in detecting and predicting seizures in epilepsy through wearable devices (WDs)[35] and in analysing rare rheumatological disorders that may take years to develop[37].

Furthermore, advanced algorithms like the Trial Pathfinder framework have been used to refine study designs by identifying common trial eligibility criteria that inadvertently exclude patients who could benefit the most from treatments. These advancements highlight the power of ML and AI in both improving patient outcomes and optimising clinical trial designs, demonstrating their role in adapting healthcare research to real-world complexities[36].

Looking ahead, RWE will continue to shape healthcare decisions and improve patient care. To expand RWE's use, stakeholders, including medicinal product and medical device developers, need to enhance their understanding of RWE analytics and implement integrated risk management processes across regions. Researchers and regulators should foster partnerships with

data analytics experts to build rapid, low-cost RWE capabilities and create high-quality, accessible RWD databases. These steps will promote innovation in RWE and improve its role in healthcare decision-making[38]&[39].

Conclusion

The importance of RWE in the healthcare industry is gaining widespread recognition. Stakeholders, including healthcare providers, researchers, policymakers, industry leaders and patients are increasingly exploring innovative approaches to leverage RWE's potential in improving patient access to safe, effective, and cost-efficient medicines. Technological advancements are expected to accelerate the adoption of RWE, although challenges persist. As awareness and acceptance of RWE continue to grow across various sectors, strategic collaborations and partnerships among all stakeholders will be essential. Engaging patients in the process ensures that their perspectives and experiences are incorporated, leading to more relevant and patient-centred data. This collective effort will facilitate the generation of actionable, broadly applicable RWE, ultimately enabling faster, more cost-effective access to treatments, including both medicinal products and medical devices, for patients.

[glossary_exclude]Further Learning Resources

- Introduction to Digital Health | EUPATI Toolbox
- Introduction to Digital Health Applications | EUPATI Toolbox
- Introduction to Benefits & Challenges of Digital Health

Processes | EUPATI Toolbox
- Introduction to Ethical, Social, and Legal Issues (ELSI) in Digital Health | EUPATI Toolbox
- Introduction to Patient Involvement in Digital Health | EUPATI Toolbox

Take now your skills to new heights by earning certification in emerging fields like Digital Health! Explore the EUPATI Open Classroom and delve into the course ‘Legal, Regulatory, and Health Technology Assessment (HTA) Concepts of Digital Health’ to deepen your understanding of the Digital Health landscape![/glossary_exclude]

[glossary_exclude]References

[1] European Medicines Agency. (2024). *Data Analysis and Real-World Interrogation Network (DARWIN EU)*. Available at: https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu (Accessed: [25.09.2024]).

[2] Cave, A., Kurz, X. and Arlett, P. (2019), Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe. Clin. Pharmacol. Ther., 106: 36-39. https://doi.org/10.1002/cpt.1426

[3] Bernard, A. *et al.* (2014) ‘Methodological choices for the clinical development of medical devices’, *Medical Devices: Evidence and Research*, 7, pp. 325–334. doi: 10.2147/MDER.S63869.

[4] Bloudek L et. Al 2021: *Are Drugs Priced in Accordance with Value? A Comparison of Value-Based and Net Prices Using Institute for Clinical and Economic Review Reports.* Value in Health (2021) 24(6) 789-794. DOI: 10.1016/j.jval.2021.01.006. (Accessed 24.11.2023 on

https://www.valueinhealthjournal.com/action/showPdf?pii=S1098-3015%2821%2900106-6).

[5] European Medicines Agency. (2022). About Us: How We Work with Big Data. Available at: https://www.ema.europa.eu/en/about-us/how-we-work/big-data (Accessed: [25.09.2024]).

[6] U.S. Food and Drug Administration. (2024). *CDER patient-focused drug development*. FDA. Retrieved November 8, 2024, from https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development

[7] Khatiwada, P., Yang, B., Lin, J.-C. and Blobel, B. (2024) 'Patient-Generated Health Data (PGHD): Understanding, Requirements, Challenges, and Existing Techniques for Data Security and Privacy', *Journal of Personalized Medicine*, 14(3), pp. 282. doi: 10.3390/jpm14030282.

[8] Dreyer, N.A. and Mack, C.D., 2023. Tactical considerations for designing real-world studies: Fit-for-purpose designs that bridge research and practice. *Pragmatic and Observational Research*, 14, pp.101-110. doi: 10.2147/POR.S396024.

[9] Dang, A., 2023. Real-world evidence: a primer. *Pharmaceutical Medicine*, 37(1), pp.25-36. doi: 10.1007/s40290-022-00456-6.

[10] Maruszczyk, K., McMullan, C., Aiyegbusi, O.L., Keeley, T., Wilson, R., Collis, P., Bottomley, C., and Calvert, M.J.(2023). Paving the way for patient centricity in real-world evidence (RWE): Qualitative interviews to identify considerations for wider implementation of patient-reported outcomes in RWE generation. *Heliyon*, e20157. https://doi.org/10.1016/j.heliyon.2023.e20157

[11] Arumugam A, Phillips LR, Moore A, Kumaran SD, Sampath KK, Migliorini F, Maffulli N, Ranganadhababu BN, Hegazy F, Botto-van Bemden A. Patient and public involvement in research: a

review of practical resources for young investigators. BMC Rheumatol. 2023 Mar 9;7(1):2. doi: 10.1186/s41927-023-00327-w.

[12] Solà-Morales O, Sigurðardóttir K, Akehurst R, Murphy LA, Mestre-Ferrandiz J, Cunningham D, de Pouvourville G. Data Governance for Real-World Data Management: A Proposal for a Checklist to Support Decision Making. Value Health. 2023 Apr;26(4S):32-42. doi: 10.1016/j.jval.2023.02.012.

[13] Lalova-Spinks T, Saesen R, Silva M, Geissler J, Shakhnenko I, Camaradou JC, Huys I. Patients' knowledge, preferences, and perspectives about data protection and data control: an exploratory survey. Front Pharmacol. 2024 Feb 20;14:1280173. doi: 10.3389/fphar.2023.1280173.

[14] Zou, K.H. and Berger, M.L. (2024). Real-world data and real-world evidence in healthcare in the United States and European Union. *Bioengineering*, 11(8), p.784. Available at: https://doi.org/10.3390/bioengineering11080784.

[15] Council for International Organizations of Medical Sciences (CIOMS) (2023). *Real-world data and real-world evidence in regulatory decision making: CIOMS Working Group report*. Geneva: CIOMS. pp. 54-56. Available at: https://cioms.ch/working-groups/real-world-data-and-real-world-evidence-in-regulatory-decision-making/(Accessed: [25.09.2024]).

[16] Oehrlein, E.M., Schoch, S., Burcu, M., Mattingly, T.J. and Perfetto, E.M. (2022). Developing patient-centred real-world evidence: Emerging methods recommendations from a consensus process. *Value in Health*, [online] 26(1), pp. 28-38. Available at: https://doi.org/10.1016/j.jval.2022.04.1738

[17] White, R. (2023). Building trust in real-world evidence (RWE): Moving transparency in RWE towards the randomized controlled trial standard. *Current Medical Research and Opinion*, 39(12), pp.1737-1741. https://doi.org/10.1080/03007995.2023.2263353.

[18] Claire, R., Cresswell, K., Avsar, T.S., Dawoud, D., Colvin, H., Chen, K., Horikoshi, Y., Kelly, N.L., Kubehl, L., Peeters, R., Sandler, S., Trejos, N., Zhu, Z., Jensen, L.G., Purcell, R., Konidaris, G., Gribbon, P., and Spring, B.(2024) *D6.2 Report on Global Regulatory Best Practices Analysis: A Scoping Review of HTA and Regulatory RWD/RWE Policy Documents*. p. 31-32. Version 2.0. Available at: https://www.iderha.org/sites/iderha/files/2024-05/D6.2%20Report%20on%20Global%20Regulatory%20Best%20Practices%20Analysis_v2.0.pdf (Accessed: [25.09.2024]).

[19] Oehrlein, E.M., Graff, J.S., Harris, J. and Perfetto, E.M. (2019) 'Patient-community perspectives on real-world evidence: enhancing engagement, understanding, and trust', *Patient*, 12(4), pp. 375-381. doi: 10.1007/s40271-019-00356-z.

[20] Berger, M. L., Sox, H., Willke, R. J., Brixner, D. L., Eichler, H. G., Goettsch, W., Madigan, D., & Makady, A.(2017). 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, 20*(8), 1003-1008. https://doi.org/10.1016/j.jval.2017.08.3013

[21] ISPOR. (2024). *Real-world evidence.* ISPOR. Retrieved [08.11.2024], from https://www.ispor.org/strategic-initiatives/real-world-evidence

[22] ISPOR. (n.d.). *Real-World Evidence Transparency Initiative*. ISPOR. https://www.ispor.org/strategic-initiatives/real-world-evidence/real-world-evidence-transparency-initiative

[23] ISPOR. (n.d.). *Real-World Evidence Registry*. ISPOR. https://www.ispor.org/strategic-initiatives/real-world-evidence/real-world-evidence-registry

[24] Canadian Agency for Drugs and Technologies in Health. (2023, May 31). *Guidance for reporting real-world evidence*. https://www.cda-amc.ca/guidance-reporting-real-world-evidence

[25] D'Andrea, E., O'Dwyer, D. N., Tcheandjieu, C., Yazdany, J., & Parikh, R. (2021). A systematic review of validated methods for identifying nonrandomized comparative effectiveness studies in the PubMed database. *Pharmacoepidemiology and Drug Safety, 30*(6), 695-702. https://doi.org/10.1002/pds.5222

[26] Dreyer, N. A., Bryant, A., & Velentgas, P. (2016). The GRACE Checklist: A Validated Assessment Tool for High Quality Observational Studies of Comparative Effectiveness. Journal of Managed Care & Specialty Pharmacy, 22(10), 1107–1113. https://doi.org/10.18553/jmcp.2016.22.10.1107

[27] Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., … & Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919. https://doi.org/10.1136/bmj.i4919

[28] National Institute for Health and Care Excellence. (2023). *Methods for real-world studies of comparative effects*. NICE. https://www.nice.org.uk/corporate/ecd9/chapter/methods-for-real-world-studies-of-comparative-effects

[29] National Institute for Health and Care Excellence (NICE). (2022). *Conduct of quantitative real-world evidence studies: Study planning*. Retrieved from https://www.nice.org.uk/corporate/ecd9/chapter/conduct-of-quantitative-real-world-evidence-studies#study-planning

[30] GETReal Institute. (n.d.). *Generating real-world evidence (RWE) for better decision-making*. GETReal Institute. https://getreal-institute.org/rwe-navigator/#generating

[31] Marra, A. R., Kobayashi, T., Suzuki, H., Alsuhaibani, M., & Shigemoto, N. (2022). Real-world effectiveness of COVID-19 vaccines: Evidence from global cohorts and clinical trials. *Frontiers in Medicine*, 9, 820544. https://doi.org/10.3389/fmed.2022.820544

[32] European Medicines Agency. (n.d.). *HMA-EMA Big Data Task Force catalogues*. EMA. Retrieved from https://catalogues.ema.europa.eu/

[33] Saesen, R., et al. (2023). Defining the role of real-world data in cancer clinical research: The position of the European Organisation for Research and Treatment of Cancer. *European Journal of Cancer, 186*, 52-61. https://doi.org/10.1016/j.ejca.2023.03.013

[34] Rudrapatna, V. A., & Butte, A. J. (2020). Opportunities and challenges in using real-world data for health care. The Journal of Clinical Investigation, 130(2), 565–574. https://doi.org/10.1172/JCI129197

[35] Beniczky, S., Karoly, P., Nurse, E., Ryvlin, P., & Cook, M. (2020). Machine learning and wearable devices of the future. Epilepsia, 62(S2). https://doi.org/10.1111/epi.16555

[36] Jreich, R., Zhang, H., Meng, Z., & Wang, F. (2024). Evaluating the robustness of an AI pathfinder application on eligibility criteria in multiple myeloma trials using real-world data and historical trials. Journal of Comparative Effectiveness Research, 13(7). https://doi.org/10.57264/cer-2023-0164

[37] Knevel, R., & Liao, K. P. (2022). From real-world electronic health record data to real-world results using artificial intelligence. Annals of the Rheumatic Diseases, annrheumdis-2022-222626. https://doi.org/10.1136/ard-2022-222626

[38] DIGITALEUROPE Executive Council for Health’s

recommendations for EU digital health policy (2024-29). (n.d.). https://cdn.digitaleurope.org/uploads/2024/02/DIGITALEUROPE-recommendations-EU-digital-health-policy-2024-29-policy-paper.pdf

[39] mHealth Data for Real World Evidence in Regulatory Decision Making An expert review report for the HMA/EMA Big Data Steering Group -2024. (2024). https://www.ema.europa.eu/en/documents/report/mhealth-data-real-world-evidence-regulatory-decision-making-expert-report-hma-ema-big-data-steering-group_en.pdf

[40] Data Saves Lives. (n.d.). Retrieved from https://datasaveslives.eu/

[41] "PICO Framework." EUnetHTA. Retrieved from .https://www.eunethta.eu/pico/.

[42] National Institute for Health and Care Excellence (NICE). (2025). *Assessing Data Suitability*. Retrieved from https://www.nice.org.uk/corporate/ecd9/chapter/assessing-data-suitability.

[43] Polak, T., Cucchi, D., Rosmalen, J., Uyl-de Groot, C. A., & Darrow, J. (2022). *Generating Evidence from Expanded Access Use of Rare Disease Medicines: Challenges and Recommendations*. Frontiers in Pharmacology, 13, 913567. https://doi.org/10.3389/fphar.2022.913567.

[44] Sedrakyan, A., Marinac-Dabic, D., Campbell, B., Aryal, S., Baird, C. E., Goodney, P., Cronenwett, J. L., Beck, A. W., Paxton, E. W., Hu, J., Brindis, R., Baskin, K., Cowley, T., Levy, J., Liebeskind, D. S., Poulose, B. K., Rardin, C. R., Resnic, F. S., Tcheng, J., Fisher, B., Viviano, C., Devlin, V., Sheldon, M., Eldrup-Jorgensen, J., Berlin, J. A., Drozda, J., Matheny, M. E., Dhruva, S. S., Feeney, T., Mitchell, K., Pappas, G. (2022). Advancing the Real-World Evidence for Medical Devices through Coordinated Registry Networks. *BMJ*

Surgery, Interventions, & Health Technologies, 4(Suppl_1), e000123. https://doi.org/10.1136/sit.2022.e000123.

[45] Stephen Gilbert, Andreia Pimenta, Ashley Stratton-Powell, Cindy Welzel, Tom Melvin, “Continuous Improvement of Digital Health Applications Linked to Real-World Performance Monitoring: Safe Moving Targets?”, *Mayo Clinic Proceedings: Digital Health*, Volume 1, Issue 3, 2023, Pages 276-287, https://doi.org/10.1016/j.mcpdig.2023.05.010.

[46] COMET Initiative. (n.d.). Core Outcome Measures in Effectiveness Trials. COMET. https://comet-initiative.org/

[47] EUPATI, “Digital Health: Introduction – Classification – Impact: 1. Introduction,” EUPATI Open Classroom, 2018, https://learning.eupati.eu/mod/page/view.php?id=975.

[48] Bolislis, W. R., et al. (2025). “Use of Real-world Data for New Drug Applications and Line Extensions.” *Clinical Therapeutics*, 42(5), 926-938. https://doi.org/10.1016/j.clinthera.2020.03.006

[49] European Medicines Agency. (n.d.). “Getting involved in EMA activities as a patient, consumer, or carer.” Available at: https://www.ema.europa.eu/en/partners-networks/patients-consumers/getting-involved-ema-activities-patient-consumer-or-career (Accessed: [28.01.2025]).

[50] TEHDAS. (n.d.). *European Health Data Space*. Retrieved from https://tehdas.eu/

[/glossary_exclude]