101191967 - GREG

Testing, improving, and co-creating Guidance and Tools for Real World Evidence Generation and Use for Decision-Making in Europe

WP1 – Evidence-based practical guidance for the use of real-world data in evidence generation for HTA and Regulatory decision-making

D1.3 Early report on RWE methodological challenges (1)

Lead contributor	Francisco Estupiñán-Romero (10. IACS) Ester Angulo-Pueyo (10. IACS) M. Soledad Isern de Val (10. IACS) Silvia Moler Zapata (10. IACS) Wendy Nieto Gutierrez (10. IACS) Berta Olano Guillén (10. IACS) Irene Sánchez Montejo (10. IACS) Javier González-Galindo (10. IACS) Tara Srouji (16. Novo Nordisk) Helene Nordahl (16. Novo Nordisk) Enrique Bernal-Delgado (10. IACS)
Other contributors	Daniel Morales (54. UEDIN) Birgit Wolf (18. BAYER) Frederique Debroucker (20. MDT) Karen Facey (2. UOXF) Leonardo Koeser (4. NICE) Anna Chaplin (29. AMGEN) Jana Boleckova (19. EW) Seamus Kent (7. EUR) Gabriela Ochoa (7. EUR)
Due date:	30/04/2026
Delivery date:	22/05/2026
Dissemination level	Public

This project is supported by the Innovative Health Initiative Joint Undertaking (IHI JU) under grant agreement No 101191967. The JU receives support from the European Union's Horizon Europe research and innovation programme and COCIR, EFPIA, Europa Bío, MedTech Europe, and Vaccines Europe.

Contents

Contents 2

Document History 4

Definitions 5

List of abbreviations 7

Executive Summary 8

1 Introduction 9

1.1 IHI-GREG Project and WP1 Objectives 9

1.2 Scope and Purpose of Deliverable 1.3 10

1.3 Foundation and Links to Prior Deliverables (D1.1) 10

2 Methods 12

2.1 Identification and Prioritisation Strategy of RWE Methodological Challenges from Scoping Search of Existing Guidelines and Recommendations 12

2.2 Collaborative Identification Strategy of RWE Methodological Challenges 12

2.2.1. Integration with Regulatory Stakeholder and HTA Expert Fora 13

2.2.2. Integration with Case Studies within the Living Library 13

2.2.3. Integration with Scope and Selection Criteria from Use Cases 14

3 Results 15

3.1 RWE Methodological Challenges Identified by the Scoping Search of the Literature 15

3.2 Early RWE Methodological Challenges in Regulatory Decision-Making 18

3.3 Early RWE Methodological Challenges in HTA 20

3.4 Early RWE Methodological Challenges from Living Library's Case Study 20

3.5 Early RWE Methodological Challenges Addressed by Use Cases 21

4 Discussion 25

4.1 Translating Methodological Challenges into Practical Guidance during GREG 25

5 Conclusion 26

Annexes 27

Annex I RWE Methodological Challenges 27

Document History

Version	Date	Description
V1.0	24/04/2026	First draft for formal review by Daniel Morales, Birgit Wolf and Frederique Debroucker.
V2.0	08/05/2026	Revised version for formal review by the GA.
V3.0	14/05/2026	Revised version with comments from GA.
V4.0	22/05/2026	Final version.

Definitions

Participants of the GREG Consortium are referred to herein according to the following codes:

1. Erasmus Universitair Medisch Centrum Rotterdam (EMC)
2. The Chancellor, Masters and Scholars of The University of Oxford (UOXF)
3. SYNAPSE Research Management Partners SL (SYNAPSE)
4. National Institute for Health and Care Excellence (NICE)
5. GETREAL Institute (GetReal)
6. Societe Europeenne De Cardiologie (ESC)
   6.1. Region Uppsala (REGION UPPSALA)
7. Erasmus Universiteit Rotterdam (EUR)
8. Stichting European Health Data and Evidence Network (EHDEN-F)
9. Stichting Eupati Foundation (EUPATI)
10. Instituto Aragones De Ciencias De La Salud (IACS)
11. University Of Galway (GALWAY)
12. The Hyve BV (THE HYVE)
13. University Of Dundee (UNIVDUN)
14. Universidade De Aveiro (UAVR)
15. Direktoratet for medisinske Produkter (NOMA)
16. Novo Nordisk A/S (NOVO)
17. Sanofi-Aventis Recherche & Developpement (SARD)
18. Bayer Aktiengesellschaft (BAYER)
19. Edwards Lifesciences Belgium BV (EW)
20. Medtronic Bakken Research Center B.V. (MDT)
21. Mölnlycke Health Care AB (MHC)
22. Institut De Recherches Internationales Servier (IRIS)
23. Pfizer Inc (PFIZER INC)
24. W. L. Gore & Associati S.R.L. (GORE S.R.L.)
    24.1. W. L. Gore et Associés S.A.R.L. (Gore S.A.R.L.)
    24.2. W. L. Gore & Associates BV (GORE BV)
25. Bristol-Myers Squibb Company Corp (BMS)
26. Boehringer Ingelheim International GmbH (BII GMBH)
27. Janssen Cilag SA (JCSPAIN)
28. GlaxoSmithKline Research & Development Limited (GSK)
29. AMGEN (AMGEN)
30. Sanofi US Services Inc. (Sanofi US Inc)
31. Sanofi-Aventis gulf F.Z.E. (SANOFI Gulf)
32. Sanofi Winthrop Industrie (SWI)
33. Sanofi Pasteur SA (Sanofi P SA)
34. Aventis Pharma Limited (Sanofi Pharma)
35. Sanofi Pasteur Limited Canada (Sanofi P L)
36. Edwards Lifesciences Sarl (EW SARL)
37. Edwards Lifesciences LLC (EW LLC)
38. Pfizer Hellas S.A. (Pfizer Hellas)
39. Pfizer AB (Pfizer AB)
40. Janssen Research & Development, LLC (JRDLLC)
41. Medical Device Business Services Inc (MDBSI)

42. Edwards Lifesciences SRL (EW SRL)
43. Edwards Lifesciences SL (EW SL)
44. Edwards Lifesciences SAS (EW SAS)
45. Edwards Lifesciences GmbH (EW GMBH)
46. Medtronic, Inc. (MDT Inc)
47. Medtronic France (MDT FR)
48. Medtronic Italia S.p.A. (MDT IT)
49. Medtronic Ltd (MDT LTD)
50. Pfizer R&D UK Limited (PFIZER R&D UK)
51. Pfizer Limited (PFIZER LTD)
52. Actelion Pharmaceuticals Ltd (ACTELION)
53. Medical Devices & Diagnostics Global Services, LLC (MDDGB)
54. The University of Edinburgh (UEDIN)

Grant Agreement (GA)	The agreement signed between the beneficiaries and the IHI for the undertaking of the GREG project (101191967)
Project	The sum of all activities carried out in the framework of the Grant Agreement.
Consortium	The GREG consortium, comprising the above-mentioned legal entities hereby referred as Participants of the GREG consortium.
Consortium Agreement (CA)	Agreement concluded amongst GREG beneficiaries for the implementation of the Grant Agreement (GA). Such an agreement shall not affect the parties' obligations to the Community and/or to one another arising from the GA.

List of abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
CAR-T	Chimeric Antigen Receptor T-cell (therapy)
CDM	Common Data Model
CHD	Congenital Heart Disease
DQA	Data Quality Assessment
ECA	External Control Arm
EF	Ejection Fraction
EHDS	European Health Data Space
EHR	Electronic Health Record
EU	European Union
GDPR	General Data Protection Regulation
HF	Heart Failure
HFmrEF	Heart Failure with mildly reduced Ejection Fraction
HFpEF	Heart Failure with preserved Ejection Fraction
HFrEF	Heart Failure with reduced Ejection Fraction
HTA	Health Technology Assessment
IHI	Innovative Health Initiative
IPD	Individual Participant Data
ITC	Indirect Treatment Comparison
JU	Joint Undertaking
KB	Knowledge Base
LLM	Large Language Model
LVEF	Left-Ventricular Ejection Fraction
MAIC	Matching-Adjusted Indirect Comparison
ML	Machine Learning
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
PPVI	Percutaneous Pulmonary Valve Implantation
RAG	Retrieval-Augmented Generation
RCT	Randomised Controlled Trial
RWD	Real-World Data
RWE	Real-World Evidence
SGLT2i	Sodium-Glucose Co-Transporter-2 inhibitors
SOP	Standard Operating Procedure
WP	Work Package

Executive Summary

This report, Deliverable 1.3, presents the early findings of the IHI-GREG project regarding the main methodological challenges currently preventing health data, collected during routing clinical practice (RWD) from being fully utilised to generate regulatory-grade RWE that can inform decisions by regulators and Health Technology Assessment (HTA) bodies. Our findings highlight a significant implementation gap, while high-level principles for RWE exist, there is a shortage of practical, contextualised, 'how-to' guidance that ensures studies are scientifically robust and reproducible.

Key findings

- Data Quality and Provenance: A lack of international consensus on how to assess if data is "fit-for-purpose" and how to maintain clinical meaning when mapping local data to global standards.
- Analytical Complexity: Challenges in replicating clinical trial results using observational data, particularly regarding bias mitigation and the precision of outcome measures.
- Sector-Specific Barriers: Unique obstacles for orphan medicines (due to small patient populations) and medical devices (where data collection often depends on proprietary manufacturer systems).
- Strategic Misalignment: The need to reconcile "top-down" regulatory requirements with "bottom-up" health system priorities to create a shared research agenda.

To address these gaps, the GREG project is developing a dynamic support ecosystem. Central to this is the GREG Knowledge Base, which feeds into an AI-powered Retrieval-Augmented Generation (RAG) system. This tool is designed to transform complex methodological literature into actionable, real-time guidance for researchers and decision-makers.

Ultimately, this report serves as a blueprint for the next phases of the project. By identifying these challenges now, GREG can focus on co-creating a functional toolkit, including standardized templates and technical benchmarks, that will improve the quality of RWE studies and improving their use in regulatory decision-making.

Next steps

The GREG project will prioritise these findings through stakeholder engagement to develop a practical toolkit to advance the use of RWE in HTA and regulatory decision-making. This includes practical guidance, recommendations, data infrastructure and tools to help researchers and decisions-makers turn complex real-world data into reliable evidence that improves access to safe and effective health technologies across Europe.

1 Introduction

1.1 IHI-GREG Project and WP1 Objectives

Health Systems in Europe produce real-world data (RWD) that can support the development and evaluation of medicines and medical devices. Despite this potential, significant barriers persist which restrict access, analysis, interpretation, and use of RWD thus limiting its impact. Furthermore, the limited adoption of existing guidelines for generating and using real-world evidence (RWE) further complicates the effective use of RWD in decision-making.

The GREG project aims to address these gaps by building upon insights from previous and ongoing key RWE initiatives. The GREG consortium will generate, pilot-test, and disseminate evidence-based guidance and tools to support the use of RWE in the development and evaluation of medicines, medical devices, and combination products.

The primary objective of WP1 is to collaboratively develop evidence-based practical guidance for generation and use of RWE. This will be done by (i) drawing on continuous assessment of existing RWE guidelines and initiatives, and (ii) collating insights from key stakeholder engagement, case study analysis, and practical use-case pilots to promote acceptance and implementation (Figure 1).

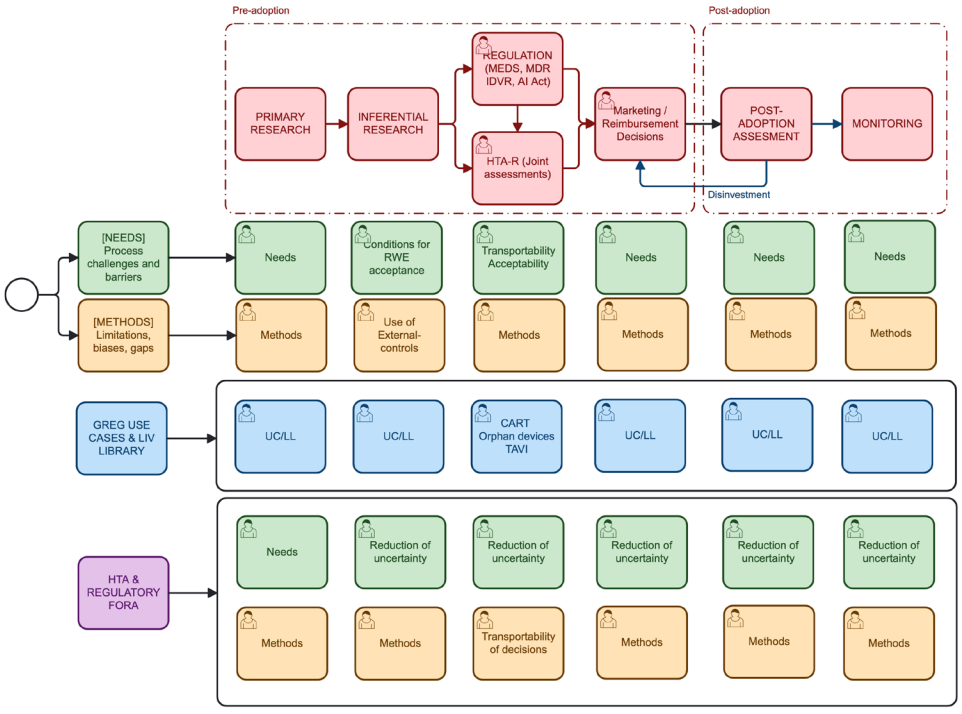

Figure 1 Conceptual framework of WP1 input streams for GREG guidance development. The schema illustrates the flow of evidence from primary sources within GREG.

The specific objective of Task 1.3 (Eliciting Data Reuse Challenges) is to synthesise insights from various sources to identify and prioritise RWE methodological challenges. These issues include data discoverability and data access procedures, data management, processing and analysis, and methodological issues and limitations in RWE generation.

1.2 Scope and Purpose of Deliverable 1.3

The scope of this deliverable is defined by Task 1.3. which primary objective is to systematically map the operational challenges associated with the secondary use of Real-World Data (RWD) within the framework of the European Health Data Space (EHDS) regulation. Deliverable 1.3 focuses on the methodological challenges for the use of RWD and RWE to support decision-making across HTA and regulatory processes.

The main purpose of this early report is to compile, prioritise, and present insights on RWE methodological challenges collated from literature (both academic and grey) scoping search, regulatory and HTA expert forums (WP2/WP3), case studies conducted within the scope of the work on creating GREG Living Library (WP3), and expectations on practical guidance provided by currently selected use cases (WP4 & WP5).

Note on Scope: Although the GREG KB includes concepts related to gaps and challenges on the interpretation of meta-analysis and Individual Participant Data (IPD) meta-analysis, which are annotated from the source materials selected as inputs, these methodologies are excluded from the present analysis, as they primarily rely on research-generated data, such as Randomised Controlled Trials (RCTs), collected in controlled experimental settings and therefore outside the report's operational definition of RWD.

1.3 Foundation and Links to Prior Deliverables (D1.1)

This deliverable can be viewed as an extension of prior Deliverable D1.1 (Early Report based on Systematic Review) and should be interpreted in the context of it. Deliverable 1.1 serves as the structural and methodological bedrock for both Deliverable 1.2 (Early Report on RWE Needs) and Deliverable 1.3 (Early Report on RWE Methodological Challenges). D1.1 describes the development of the initial GREG Knowledge Base (KB) and the core GREG ontology (v.0.0.1). This is achieved through a comprehensive scoping search of existing RWE guidelines and recommendations. D1.2 and D1.3 utilise the updated GREG KB after completing the GREG ontology validation (v.0.0.2). D1.2 and D1.3 leverage GREG RAG capabilities to identify challenges, limitations, and gaps in generating RWE from the literature by semantically annotating key related concepts, grouped into distinct dimensions and categories.

The GREG KB and Ontology (v.0.0.2) includes over 25,000 concepts. These are contextualised on 481 dimensions and explicitly categorised under 'Limitations, Gaps, and Challenges', 'Methodology and Processes', and 8 other categories. The concepts were semantically annotated, extracted, and grouped from 652 processed documents (see D1.1). The GREG KB is set to act as a continuously updated *single source of truth* that GREG can rely on to accurately map stakeholder needs and methodological barriers and retrieve practical guidance on the use of RWD and RWE in HTA. The GREG KB is part of the Retrieval-Augmented Generation (RAG) system set up by WP1. It can be directly queried using Cypher language [https://docs.falkordb.com/cypher/] on FalkorDB [https://www.falkordb.com/] as a graph database. Alternatively, it can be accessed via a smart chatbot interface that enables inquiry-response (Q&A) interactions, powered by a generative Large Language Model (LLM).

Future versions of the GREG KB and RAG systems are expected to continuously synthesise and retrieve:

- Scoping searches of existing literature and global guidelines,
- Examples of use of RWE in regulatory and HTA submissions as catalogued by GREG Living Library,
- Methodological lessons and data reuse challenges and solutions extracted from GREG use cases for medicines and devices,
- Expert opinions elicited from Regulatory and HTA expert forums use of RWE.

The GREG KB (v0.0.2) focused on processing prioritised documents providing guidance published by decision-making agents, or international initiatives on RWE between 2015 and May 2025 (*see D1.1 to check the search strategy, selection and prioritisation criteria*).

2 Methods

Mapping the current and emerging RWE methodological challenges requires deploying a mixed-methods approach, combining a scoping search with advanced AI-driven document processing, with insights from the work done in other WPs engaging regulatory stakeholders (WP2) and HTA body experts (WP3), from regulatory and HTA bodies' documented decisions (WP3), and from the scope and the practical challenges posed by selected use cases (WP4 & WP5) based on the reuse of RWD (WP6).

2.1 Identification and Prioritisation Strategy of RWE Methodological Challenges from Scoping Search of Existing Guidelines and Recommendations

RWE methodological challenges were identified by querying the GREG KB (v0.0.2) to retrieve concepts within *methodological challenges* dimension, under the '*Limitations and Challenges*' category of the GREG ontology (v.0.0.2).

The GREG Knowledge Base and Ontology (v.0.0.2) includes 25,977 concepts, contextualised on 481 dimensions and explicitly categorised under 10 categories; semantically annotated, extracted, and aggregated from 652 processed documents. A total of 6,652 concepts are included in those dimensions under the '*Limitations and Challenges*' category. To end up with a usable set of RWE methodological challenges to focus on for D1.2 and D1.3, concepts were selected in a stepped process following several prioritisation criteria:

1) Concurrency: Concepts selected when they appeared within the same context in more than one source (i.e., document or guidance).
2) Frequency or density of evidence: Concepts ranked based on their frequency of appearance within the selected dimensions (i.e., gaps, limitations, barriers, and recommendations).
3) Relevance to the assessment paradigms: Concept specifically linked to HTA or regulatory workflows, methodologies and processes, thus reflecting specific operational hurdles faced by sponsors or decision-makers or whether they were general to using of RWD/RWE in this context.
4) Validation and stakeholder alignment: Top ranked concepts in the final selection were compared with RWE methodological challenges mentioned during stakeholder interactions (regulatory and HTA experts) and with those identified by case study analysis (living library), including the scope and challenges addressed by the pilot use cases.

2.2 Collaborative Identification Strategy of RWE Methodological Challenges

While a continuous scoping literature search is foundational to the RWE challenges identification strategy, it may not fully capture the nuanced operational realities faced by decision-makers. As a result, WP1 relies heavily on the targeted stakeholder engagement done throughout the project by other WPs and the structured qualitative feedback that is produced. This engagement ensures outputs reflect practical emerging demands, helps prioritise methodological challenges as well as critical appraisal of potential knowledge gaps and unmet needs.

This identification strategy is built on collaboration with WP2 (Regulatory) and WP3 (HTA) and benefits from the structured qualitative analysis of outputs (i.e., summaries and reports) from Regulatory and HTA expert forums. Crucially, frequent stakeholder engagement throughout the project can act as a real-world validation mechanism by cross-referencing the theoretical gaps, limitations and barriers identified in the literature with those elicited by relevant stakeholders, and by ensuring that the RWE methodological challenges mapped are accurate, highly contextualised and directly actionable for the subsequent co-creation of the GREG practical guidance.

This mixed methods approach is operationalized through the integration of insights from the literature with findings derived from structured qualitative and narrative analyses of various information streams. These streams include stakeholder engagement, pilot use cases, and case studies on the use of RWD/RWE in HTA and regulatory decision-making, the latter completed as part of the Living Library. This approach ensures that WP1's identification of RWE needs remains iterative and responsive to the evolving healthcare and regulatory landscape throughout the GREG project. The analysis involves a structured interpretation of insights from each stream, categorized using the GREG ontology (v0.0.2) as a reference, followed by a narrative synthesis both within their specific contexts and the overarching scope of the GREG project.

To ensure these collaborative insights are effectively harnessed, all GREG outputs are expected to be iteratively integrated within the GREG KB, summarizing insights from all WPs into a centralized, accessible evidence base.

2.2.1 Integration with Regulatory Stakeholder and HTA Expert Fora

This stream focuses on formal collaboration with WP2 (Task 2.1: Regulatory Stakeholder Forum) and WP3 (Task 3.2: HTA Expert Forum). Leveraging output from the forums, WP1 gathers high-level institutional insights about specific RWE methodological challenges associated with using RWD to generate regulatory-grade evidence. The structured qualitative analysis of meeting summaries enables identification of specific challenges in the technical application of RWD/RWE.

This deliverable includes insights produced from the structured qualitative and narrative analysis of reports from:

- First Regulatory Stakeholders Forum meeting on April 2nd 2026 (WP2), and
- First HTA Experts Forum meeting on February 23th 2026 (WP3).

2.2.2 Integration with Case Studies within the Living Library

This stream uses the practical evidence generated by the Living Library (WP3), particularly on the deep-dive case studies. By analysing their results, WP1 can extract real-world insights into the methodological challenges and data quality issues encountered during the implementation of RWE in HTA and regulatory decision-making. The deep-dive case studies are designed to reverse-engineer actual HTA submissions to extract empirical actionable insights on methodological barriers when using RWD/RWE.

Deep-dive case studies conducted as part of the Living Library can span the entire product lifecycle (from pre-authorisation to post-marketing evidence generation) and can cover medicines, medical devices, and drug-device combinations across various disease areas. Analysing RWE use in those cases, contextualised by the positive or negative result of the assessment in each regulatory environment, can clarify the varying evidence standards demanded by different EU HTA bodies and highlight practical barriers to use RWE. More importantly, this analysis provides the necessary information to determine which specific analytical or procedural issues potentially compromising the acceptability of the evidence. Consequently, it helps identify the underlying factors that influence HTA bodies' willingness to accept observational data, focusing specifically on the intersection between the fitness-for-purpose of the underlying RWD and the methodological robustness of the generated RWE.

This deliverable includes insights produced by the structured qualitative and narrative analysis of the outputs of the focused research study on the evaluation of Chimeric Antigen Receptor T-cell (CAR-T) therapies included in deliverable D3.2 Living Library of HTA Use Cases (WP3).

2.2.3 Integration with Scope and Selection Criteria from Use Cases

The collaborative strategy also incorporates feedback from the scope, methodology, and technical requirements of the pilot use cases selected by WP4 (Medicines) and WP5 (Medical Devices and Drug-device combinations). Selection of use cases was based on the reuse of RWD data accessed from Data Partners within the GREG consortium (WP6). These use cases can serve as practical benchmarks for testing or contextualising the future guidance developed by GREG. Their contribution to identification of RWE methodological challenges lays on the capacity to provide direct practical insight into several high-priority methodological areas (e.g., data provenance, fitness-for-purpose, interoperability, standardisation and other technical challenges, bias mitigation and confounding). GREG can check where existing guidelines fail to provide actionable instructions by applying specific methodological frameworks to relevant research questions using available RWD.

Finally, the results from these use cases will be iteratively fed back to the GREG KB, ensuring that the challenges identified and prioritised in this deliverable are part of an evolving repository that reflects the operational realities of generating evidence robust enough for regulatory and HTA decision-making.

This deliverable includes insights produced by the structured qualitative and narrative analysis of the protocols of the use cases selected for piloting in:

- Heart Failure with preserved, moderately reduced and reduced ejection fraction (WP4), and
- Transcatheter Pulmonary Valve Bio Prosthesis – Methods for Incidence and Prevalence Estimation in Orphan Medical Device Designation with Electronic Health Records (WP5).

3 Results

3.1 RWE Methodological Challenges Identified by the Scoping Search of the Literature

Collated insights from GREG Knowledge Base and GREG Ontology v0.0.2.

From an initial pool of 6,652 concepts across those dimensions under the '*Limitations and Challenges*' category, a total number of 417 were identified as methodological challenges by retrieving all concepts to methodological issues, gaps, deficiencies and challenges. This framework assessed concepts based on their (i) concurrence across multiple sources, (ii) frequency of appearance and (iii) relevance to specific HTA or regulatory workflows. Ultimately, ranking the RWE methodological challenges by their prevalence within the KB enables GREG to focus on the key challenges or issues (classified into 6 main areas) and ensures alignment with relevant literature.

Areas	Challenges	Frequency (N)
GENERAL	STUDY DESIGNS COMMON METHODOLOGIES TOOLS	10
PROCEDURES	METHODOLOGICAL APPROACHES TOP-DOWN BOTTOM-UP	12
STANDARDS	FRAMEWORKS STANDARDS CONCEPTS	26
TECHNIQUES	DATA ANALYTICS MACHINE LEARNING REGULARIZATION METHODS SIGNAL DETECTION METHODS SYNTHETIC CONTROL METHODS CAUSAL INFERENCE METHODS BIAS & UNCERTAINTY ANALYSIS METHODS	21
MODELS	BENEFIT-RISK ASSESSMENT METHODOLOGIES BENEFIT-RISK METHODOLOGY APPRAISAL OPPORTUNITY-COST METHOD PROXY GOOD METHOD	19
DEFICIENCIES	LACK OF STANDARDIZED METHODS LACK OF UNIVERSALLY ACCEPTED METHODOLOGICAL STANDARDS METHODOLOGICAL CHALLENGES METHODOLOGICAL GAPS	329

Areas	Challenges	Frequency (N)
	COMMON METHODOLOGICAL ISSUES	
	LIMITATIONS	
	INCORRECT ADJUSTING METHODS	
	LACK OF SENSITIVITY ANALYSIS	

High-level RWE methodological challenges included a lack of frameworks, overarching protocols, systematic processes or international consensus in various areas:

- **Inadequate frameworks for RWE assessment:** Present frameworks prioritise procedural design and basic analytical principles for scientific validity. However, technical guidance remains limited to fundamental research components—such as randomisation, control selection, and error-rate management. Insufficient attention is paid to advanced bias control, outcome precision, and ensuring data are "fit-for-purpose" through rigorous validation.
- **Poorly defined data acquisition protocols:** Overarching protocols for data acquisition are still underdeveloped. This results in vague specifications regarding data provenance and identification, as well as a lack of rigorous inclusion/exclusion criteria based on thematic or temporal parameters.
- **Lack of meta-methodological frameworks:** There is no clear system for selecting candidate methodologies or comparing the outputs of competing approaches. Evaluation parameters—such as complexity, interpretability, and accessibility—are frequently undefined, making it difficult to justify one methodological choice over another.
- **Absence of standardized RWD evaluation frameworks:** Current methodologies lack a structured approach to evaluate and benchmark RWD analysis processes. To bridge this gap, the field requires established core principles (e.g., data integrity and methodological validity) alongside operational metrics (e.g., scoring frameworks to evaluate data utility and sensitivity). Developing a unified evaluation framework is critical for aligning fields with different objectives, such as pharmacoepidemiology (safety-focused) and HTA (value-focused), facilitating cross-sector acceptance of RWE.
- **Methodological gaps in Orphan Medicines and Devices:** Current RWD analysis methods fail to adequately manage clinical uncertainties inherent in rare diseases. There is a need for specialized RWE study designs tailored for rare and ultra-rare diseases. Generating robust RWE requires mechanisms for cross-border data linkage, alongside structured frameworks for multi-stakeholder engagement and specific methodological approaches.
- **Reproducibility of data extraction and transformation outputs and database benchmarking:** There is no formal approach to testing the reproducibility and reliability of data extraction and transformation methods. Furthermore, the field lacks benchmarks to compare similarities and discrepancies across health databases with varying technical specifications and architectures.
- **Fragmentation in the application of existing instruments:** While numerous tools exist (e.g., REQueST, STROBE, GRACE, STRATOS, SPACE, TARGET), they are not integrated into a coherent methodological guide.

There is a need to align these instruments with the specific requirements of HTA and regulatory decision-making throughout the health technology life cycle.

- **Lack of standardised Data Quality Assessment (DQA):** No international consensus exists regarding data governance or DQA dimensions in healthcare. Assessment focus remains fragmented among basic aspects: foundational (provenance and documentation), intrinsic quality (fitness-for-use), and question-specific requirements (fitness-for-purpose).

On top of these general issues, there is an ongoing discussion regarding top-down versus bottom-up approaches to RWE generation. The top-down model involves HTA and regulatory bodies creating an explicit demand for RWE in specific use cases. Conversely, the bottom-up model focuses on health systems and researchers generating evidence based on their own priorities, following a directive to embed implementation science within research.

This debate signals a broader need for a systematic approach to organising research based on the reuse of RWD. At a policy level, it highlights the need to align the interests of regulatory bodies with those of health systems to create a complementary research agenda. Methodologically, it presents the challenge of systematising RWD research so that it simultaneously informs decision-making at both the policy and health services levels.

In addition to those more general challenges, specific methodological challenges arise from:

- **Lack of consensus or standardised use of existing concepts:** There is an abundance of polysemy and context-dependent jargon in both RWD management and methodology fields. For instance, terms like 'data provenance', 'real-world evidence', and 'validation' are used inconsistently across regulatory and clinical research settings. A similar case, happens with methodologies and data analysis techniques, such as 'causal inference', 'machine learning', or 'AI', that get a different meaning assigned depending on the area of research;
- **Excess of narrowly focused frameworks:** Existing guidelines often operate in silos, focusing exclusively on specific niches, such as privacy-preserving technologies, specific therapeutic areas, or isolated stages of the data life cycle (i.e., collection vs. analysis). This fragmentation prevents the development of an end-to-end methodological 'map' that researchers can follow from the initial data extraction in a health system (bottom-up) to a final submission for HTA reimbursement (top-down), and
- **Poorly adopted standards:** Despite the availability of high-quality technical standards, their implementation remains inconsistent due to the high resource burden of data mapping (e.g., to OMOP Common Data Models) and the lack of incentives for health systems to maintain research-grade data documentation. This leads to a 'legacy debt' where RWE is generated on non-interoperable infrastructures that cannot be easily replicated or audited by external regulatory bodies.

Finally, literature analysis revealed the lack of specific practical guidance, instructions, or even tutorials on the specific application of certain data analysis techniques or models for producing RWE for HTA or regulatory decision-making, particularly:

- Machine Learning (ML) and applied AI methods,
- Variable regularisation methods,

- Signal detection and quality performance control methods,
- Synthetic control methods,
- Causal inference methods,
- Bias & uncertainty analysis and assessment methods,
- Benefit-risk assessment or appraisal methods,
- Opportunity-cost modelling methods, and
- Proxy-good modelling methods, among others, which also suffer from lack of cost or price data or lack of accessibility.

The convergence of these systemic and technical challenges creates a significant methodological bottleneck and an a priori mistrust in RWE, also driven by the unrealised promises of RWD. The tension between top-down regulatory demands and bottom-up health systems priorities, compounded by inconsistent terminology and fragmented frameworks, results what seems to be pervasive methodological deficiencies in current RWE studies. These deficiencies, ranging from poor study specification to lack of data provenance or reproducible protocols, undermine the reliability of evidence use in critical HTA and regulatory decision-making pushing its role to those areas or technologies for which there is no other evidence available and limiting its implementation in post-market evaluation for reassessment.

3.2 Early RWE Methodological Challenges in Regulatory Decision-Making

Insights on RWE methodological challenges commented by regulatory stakeholder forum experts in their first meeting were presented in the context of current RWE needs and perceived gaps. These could be classified into the following broad categories:

- Cross-cutting data and guidance challenges:
  - Data access and data quality assessment
  - Implementation gap from high-level guidance and lack of common standards
  - Lack of feasibility assessment guidance
  - Low overall quality of RWE studies supporting regulatory decision-making, leading to uncertainty in their utility and concerns regarding their limitations
- Advanced analytical challenges:
  - RCT emulation and replication: Difficulty of application of the Target Trial emulation framework when 'time zero' (index date) is poorly defined in observational records, leading to immortal time bias, in addition to potential unmeasured confounding, and the difficulty to clearly establish exposure to a medicine or medical device in routine clinical practice.
  - Effectiveness and indication extensions: Methodologies for proving effectiveness in new populations often struggle with *protopathic bias* (i.e., reverse causation) and the lack of a clear counterfactual in real-world clinical practice.
  - Federated analysis or data pooling to overcome small sample size issues: Technical barriers include the 'black box' nature of some federated algorithms where centralised meta-analysis is impossible, alongside legal complexities regarding GDPR-compliant cross-border data flows.
- Challenges specific to Medicines:

- ☐ Clinical studies with external control arms (ECA): The primary challenge is ensuring comparability (balance) between the intervention trial arm and the ECA, particularly regarding differences in baseline standard of care and ancillary treatments.
- ☐ Clinical studies on populations excluded from RCTs: Methodological tools are lacking to handle the increased noise and comorbidities found in paediatric, geriatric, or pregnant populations, where physiological changes can affect drug metabolism in ways not captured by standard RWD or in rare diseases were sample sizes are small, there is no comparators available and evolution of the disease is highly dependent on healthcare and other external determinants.
- ☐ Lack of information on the impact of data mapping to common data models (i.e., losses in semantic interoperability due to conformance with common data models): Mapping to Common Data Models (CDMs) like OMOP often can result in information loss where local clinical nuances are forced into a standardised format that may not perfectly fit the original clinical intent.

🕮 Challenges specific to medical devices (and combination products):

- ☐ Data access: Unlike medicines, devices are often not captured by prescription registries nor utilisation registries. Access to relevant information is also limited by the limited use of universal medical device identifiers in claims data, routine billing or clinical notes. In addition, information on device's version in use is effectively missing in all cases except for failure or safety reports.
- ☐ Lack of examples from regulatory cases for medical devices and particularly for combination products: There is a general absence of examples of successful device-related RWE submissions, making manufacturers potentially hesitant to invest in innovative study designs.
- ☐ Study designs without controls
  - Lack of plans for post market clinical follow up, including plans for data collection/data access and analysis: Many devices lack a predefined strategy for longitudinal tracking, making it nearly impossible to assess long-term safety or 'device drift'.
  - Lack of studies characterising medical device indications within target population: There is a lack of real-world mapping defining exactly how a device is used in the health services versus how the device is supposed to be used (i.e., Instructions for Use).
- ☐ Post-market generation and device-generated data
  - Data collection depending on manufacturers: Data is often trapped in proprietary silos that are inaccessible to independent regulatory decision-makers or health system researchers.
  - Lack of structured health information on medical devices utilization: Health records rarely document the procedural context (e.g., surgeon experience, specific surgical technique or anatomical information on the patient undergoing a procedure) which heavily influences device performance.
  - Lack of structured information of the versions of devices being used in health care: The rapid interaction of hardware and software, more so in devices supporting the deployment of ML/AI-based algorithms, means that by the time an RWE study is completed, the device version analysed may already be obsolete, setting the need for periodic reassessment.

3.3 Early RWE Methodological Challenges in HTA

Collated insights from the WP3 HTA Expert Forum meetings detailed HTA evidence standards, post-marketing evidence generation challenges, and lifecycle reassessment demands.

The HTA expert forum meeting highlighted several specific methodological challenges requiring guidance or tools development regarding RWE:

- **Quantifying the risk of bias and uncertainty:** Because RWE is never completely conclusive, a major methodological challenge is systematically understanding and quantifying the risk of bias and residual uncertainty. The forum suggested that approaches like utilizing a Bayesian method for quantification could be proposed to comprehensively consider how RWE characteristics impact uncertainty. Furthermore, experts stressed that this risk of bias needs to be explicitly linked to frameworks that support decisions, rather than solely focusing on how decisions are made.
- **Mapping to Common Data Models (CDMs):** With the widespread use of federated data analytics, one expert identified a need for more research to understand the potential loss of data quality when mapping federated data to a CDM, as well as determining which CDM is best to use.
- **Further exploration of new analytical methods in HTA:** One expert noted that new quantitative transportability methods, which are used to estimate treatment effects from RWE generated in other jurisdictions, need further exploration because they are not currently widely used in HTA. Other expert pointed to the need to explore and assess implications of the use of federated analyses.
- **Using RWD to mitigate cost-effectiveness uncertainties:**
    - Developing new pricing, cost-effectiveness, and reimbursement models.
    - Addressing other perspectives (i.e., societal perspective) in economic modelling.

In addition, HTA experts commented on the demand for increasing or enhancing HTA bodies capabilities by developing specific training and capacity building programs on the RWD analysis frameworks and methodological guidance on RWE generation, thus:

- **Developing assessor capabilities:** Addressing these complex methodological challenges, particularly around uncertainty analysis, requires the active development of new methods alongside the enhancement of HTA assessor capabilities to evaluate RWE submissions. Experts identified building this capability around uncertainty analysis as a highly valuable area for collaboration among HTA bodies and regulators.

3.4 Early RWE Methodological Challenges from Living Library's Case Study

Preliminary insights from the ongoing case study on CAR-T therapy conducted as part of the work for the Living Library, reveal several recurring methodological challenges and limitations when evaluating RWE. The main methodological challenges identified include:

- **Uncertainties in Indirect Treatment Comparisons (ITCs**): Because many CAR-T therapies are approved based on single-arm trials, sponsors frequently use ITCs, such as Matching-Adjusted Indirect Comparisons (MAICs) or propensity-score matching, to compare trial data with external or historical RWE cohorts. HTA bodies frequently reject these comparisons or consider them highly uncertain because of the **lack of direct comparative evidence and difficulties in ensuring the populations are truly comparable**.
- **Confounding and Residual Bias:** HTA bodies consistently point out that RWE studies suffer from residual confounding because **it is impossible to adjust for all relevant confounding factors**. For example, in

some propensity score models, the failure to include key prognostic factors (such as ECOG status, sex, or disease stage) does not hold the crucial assumption that there is no unmeasured confounding.

- **Small Sample Sizes:** When matching trial patients to RWE cohorts, the resulting effective sample sizes are often very small, which limits statistical power and the confidence HTA bodies have in the comparative effect estimators. In one noted instance in the CAR-T in-depth case study, the sample of patients retained for an indirect comparison represented only 20% of the overall real-world cohort.
- **Lack of Generalisability (External Validity):** HTA committees frequently question whether the results from tightly controlled pivotal trials can be transposed to real-world practice. Clinical trials often have strict inclusion criteria limited to patients in good condition with a low risk of complications, which does not accurately reflect the broader target population. Additionally, differences in standard treatments and follow-up care across different countries (e.g., Dutch clinical practice versus study cohorts) limit the relevance of the data.
- **Complexity and Lack of Transparency:** Some HTA bodies criticise the methods used to select and adjust RWE data for being highly complex or incomprehensible. If the procedure for identifying and matching the external control cohort is unclear, it undermines the reliability of the entire analysis.
- **Absence of Sensitivity Analyses:** HTA bodies note that when sponsors fail to include sensitivity analyses in their RWE submissions, it becomes impossible for evaluators to quantify the potential magnitude of the residual bias present in the data.

3.5 Early RWE Methodological Challenges Addressed by Use Cases

The identification of RWE methodological challenges within the GREG project is deeply rooted in the empirical findings of ongoing use cases (WP4 & WP5) and from the work done to facilitate access to RWD within the scope of the GREG consortium (WP6). This subsection collates early insights into the methodological, technical, and structural barriers encountered during evidence generation.

These pilot studies are meant to serve as a 'stress test' for current RWE frameworks, transitioning from high-level conceptual challenges to the granular, operational difficulties and methodological challenges faced by data scientists and researchers in generating regulatory-ready RWE.

The insights derived from the uses cases, currently ongoing, can be categorised into two primary dimensions: (i) procedural execution, and (ii) the definition of the 'feasibility frontier' through the application of standard selection criteria and data feasibility assessments.

- **Procedural execution:**
  - **Data Provenance**: The development of the proposed use cases is supported by the GREG data catalogue (WP6) providing standardised information (i.e., metadata) on the data available to the GREG consortium. It sets a baseline requirement and a challenge to other research networks or RWE generation initiatives to adopt data discoverability and data management standard operational procedures produced and tested within the GREG project.
  - **Fitness-for-Purpose:** Evaluating the quality of the data, available to the GREG consortium, in terms of their fitness for specific purposes, these use cases identify potential systemic discrepancies between actual routine clinical data and documentation, and the data structures required for robust research. This includes quantifying missingness, and potential recording bias for variables and information essential to HTA and regulatory decision-making.
  - **Interoperability and Standardisation:** The leverage of the OMOP CDM as a de facto standard, the implementation of standard operational procedures for the landscape analysis and the feasibility analysis covering both the data accessibility and the methodological robustness, highlights the technical challenge in legal, organizational, semantic, and technological interoperability. Specifically, it surfaces

the tension between data standardisation and the preservation of clinical nuance during the transformation of local electronic health records (EHR) into common formats and schemas.

- **Definition of a potential 'feasibility frontier':**
  Beyond the active analysis, the criteria used to select these use cases constitute valuable insight on RWE challenges. Documenting why specific datasets or clinical scenarios were deemed unsuitable, considering factors such as lack of data or the required data granularity, insufficient longitudinally or timelines of the available data, or legal and organisational constraints to access data, use cases development help mapping the current 'feasibility frontier' of RWE.

Each use case is set to assess GREG's capability to test and produce evidence on a specific evidence gap (for Medicines, Medical devices, or both) that address specific methodological challenges. A summary of each use case is provided in the table below with specific reference to the RWE gap or challenge being addressed.

Use Case 1 (WP4 – Medicines): Chronic Heart Failure by Ejection Fraction Phenotypes

Scope: This study focuses on patients with incident chronic heart failure (HF), stratifying them by left-ventricular ejection fraction (LVEF) phenotypes: reduced (HFrEF), mildly reduced (HFmrEF), and preserved ejection fraction (HFpEF).

Aim and Primary Objectives: The overall aim is to estimate the incidence of HF and to characterise patient populations, treatment pathways, and outcomes across the different ejection fraction (EF) subtypes. The primary objectives are to:

1. Estimate the incidence of HF overall and by EF subtype.
2. Characterise patients with incident HF by demographics, comorbidities, healthcare utilisation, survival, and major cardiovascular outcomes, overall and by EF subtype.
3. Characterise real-world patients with incident HF who initiate treatment with Sodium-Glucose Co-Transporter-2 Inhibitors (SGLT2i) to compare their profiles and outcomes with those in registry-based clinical trials.

Methodology: Retrospective, observational cohort study utilising RWD from different databases across EU countries and USA. All databases are mapped to OMOP CDM, and the study is intended to use a federated analysis approach, running analytical code locally using OHDSI and DARWIN EU R packages without centralising patient-level data.

RWE methodological challenges addressed:

- **Challenge: Handling missing clinical data in routine datasets**. Most health information systems based on EHR do not routinely capture EF measurements as structured data, resulting in a high rate of missing data. In addition, specific EF-based diagnostic codes are often unmapped or only available in unstructured data (i.e., medical notes or diagnosis images records).
    - **Methodological Lesson/Guidance**: The study will provide lessons on extracting and mapping critical phenotypic data from unstructured text. The protocol aims to provide guidance on managing missingness and understanding the limitations of health databases with large volumes of unrecorded clinical parameters.
- **Challenge**: **Proving trial representativeness and transportability:** HTA bodies frequently face uncertainty regarding whether the treatment effects seen in highly selected trial populations (e.g., for SGLT2 inhibitors) are transportable to real-world patients across different HF phenotypes.
    - **Methodological Lesson/Guidance**: By comparing real-world treatment initiators with trial cohorts, the use case can yield methodological guidance on how to evaluate and quantify the real-world applicability of trial results, establishing a framework to satisfy uncertainties regarding treatment effectiveness in broader EU populations.

Use Case II (WP5 – Medical devices): Orphan Medical Device Designation (Transcatheter Pulmonary Valve Bioprothesis)

Scope: This study focuses on patients with congenital heart diseases (CHD) involving right ventricular outflow tract (RVOT) dysfunction (such as Tetralogy of Fallot or common arterial trunk) who have undergone RVOT reconstruction and subsequently developed residual lesions requiring reintervention via percutaneous pulmonary valve implantation (PPVI) of a bioprothesis. **Aim and Primary Objectives:** The primary aim is to compare methods for estimating incidence and prevalence of vary rare conditions using RWD to support orphan medical device designation. The primary objectives are to:

1. Estimate the incidence and prevalence of the underlying CHD population, the post-RVOT reconstruction population, and the sub-population that is ultimately eligible for PPVI.
2. Compare the RWD-based incidence and prevalence estimations with prior literature and national statistics.
3. Characterise the populations in each cohort and their drug use/treatment patterns.
4. Compare methods for combining evidence (meta-analysis, triangulation, etc.) across multiple OMOP-mapped databases to extrapolate incidence and prevalence results to the EU-27 population-level to respond to the orphan medical device designation criteria.

Methodology: Retrospective, population-level descriptive analysis, observational cohort study utilising RWD from different databases across EU countries and USA. All databases are mapped to OMOP CDM, and the study is intended to use a federated analysis approach, running analytical code locally using OHDSI and DARWIN EU R packages without centralising patient-level data. It incorporates the requirement for structured triangulation or meta-analysis for direct rate scaling of standardisation to extrapolate small event counts from various contexts to a broad EU level.

RWE methodological challenges addressed:

- **Challenge: Estimating incidence of highly specific subpopulations (small event counts):** The target population is people with CHD, who had previous RVOT that subsequently developed specific progressive dysfunctions requiring reintervention. Such granular data is rarely captured in routine national statistics, and filtering for it in RWD results in extremely small event counts.
    - **Methodological Lesson/Guidance:** Step-by-step guidance on denominator selection (i.e., phenotyping) and cohort funnelling.
- **Challenge: Extrapolating regional RWD to and EU-27 population level:** Confidently scaling up local or regional incidence rates from a handful health data partners to the entire EU population is fraught with uncertainty and relies on potentially outdated or unstratified sociodemographic population statistics.
    - **Methodological Lesson/Guidance:** Guidance on compared methods for EU-wide extrapolation.
- **Challenge: Triangulating highly heterogeneous evidence:** The use case will evaluate multiple candidate pooling methods, quantifying between-database and method variability.
- **Challenge: Addressing unmeasured confounding in rare diseases.** The use case implementation will highlight the limitations of using RWD for regulatory thresholds.

4 Discussion

4.1 Translating Methodological Challenges into Practical Guidance during GREG

The translation of the RWE methodological challenges identified in this early report into actionable guidance within the GREG project follows a targeted iterative pathway. By mapping high-frequency challenges in different areas (e.g., technical data quality assessments, lack of standardised terminology, methodology implementation, complexities of federated analytical replication), the project can prioritise the development of standardised operating procedures (SOP) and technical benchmarks that overcome the current divide between guidance and practice. Specifically, the qualitative nuances extracted from the limitations and challenges can serve as the first step towards a structured blueprint for the practical frameworks in later work packages, ensuring that future recommendations and guidance are grounded and offer actual settings, metrics, thresholds, instruments, and tools to guide the use of RWD/RWE to impact HTA and regulatory decision-making at the scale required by the current pace of advancements in health technology.

Ultimately, this evidence-driven prioritisation and gap analysis allows GREG to focus its resources on creating modular scalable tools that address the most critical barriers to RWE adoption, transforming the raw knowledge base concepts into a cohesive roadmap for regulatory and HTA excellence.

5 Conclusion

The identification of diverse methodological challenges underscores the need for an iterative stakeholder-driven approach within the GREG project. While the initial landscape mapping in this early report has identified the main challenges in using RWE for decision-making, the path forward requires prioritising efforts to transform these insights into actionable guidance. The transition from a static understanding to a dynamic support ecosystem will be driven by several strategic actions.

First, WP1 is expected to update and evolve this report through multiple consecutive milestones:

- D1.6 & D1.10 (Early Report of RWE Methodological Challenges II & III): These reports will iteratively update the initial findings by incorporating new methodological insights gained from the evaluation of existing guidelines and the subsequent pilot-testing of GREG guidance.
- D1.7 & D1.11 (RWE Adoption Challenges Report I & II): These deliverables will focus on describing and analysing the barriers that HTA and regulatory bodies face when implementing RWE, ensuring the final guidance is practically applicable to these institutions.
- D1.12 (Data Reuse Report): This report will synthesise the specific challenges related to accessing and reusing sensitive health data for regulatory purposes within the context of the European Health Data Space (EHDS).

Second, WP1 acts as a central hub of the project by providing the context, awareness and requirement analysis to synthesise empirical lessons, recommendations, and practical guidance produced within other WPs into a harmonised and structured framework to accelerate the use of RWE. Its success is tethered to the technical maturation and continuous update of the GREG RAG system:

- **GREG KB enrichment:** Insights from the Living Library and the HTA/Regulatory forums will be integrated and semantically annotated into the GREG KB, ensuring the RAG system does not just rely on published literature, but also incorporates consensus-built recommendations produced within the project.
- **Iterative prompt engineering and fine-tuning:** As WP4 and WP5 pilots encounter and potentially overcome technical and methodological challenges, these lessons learned and the tools tested and proved effective will be used to refine the RAG system's response logic, facilitating proper guidance grounded in the practical reality of analysing RWD as response to users' querying.

This deliverable and its future versions will support WPI's cross-work package feedback loops, ensuring ongoing alignment with WP2, WP3, and WP6. This approach will help map tools, procedures, and templates to the high-priority gaps identified here, resulting in a functional toolkit customised to the EU regulatory and HTA landscape. Consequently, addressing these methodological challenges through appropriate practical guidance, and promoting the adoption of standardised approaches to correct existing study deficiencies and improving the quality and relevance of RWE in Europe.

Annexes

Annex I RWE Methodological Challenges

It can be accessed as an Excel file: Annex I_extracted_prioritized_methodological_challenges_for_synthesis.xlsx