Analysis and interpretation of real-world data: a 5-year outlook

Patrick Ryan, PhD
Janssen Research and Development
Columbia University Irving Medical Center
Disclosure

Opinions are my own, and do not necessarily reflect those of Janssen R&D, Columbia University or OHDSI community
Real-World Evidence In EU Medicines Regulation: Enabling Use and Establishing Value

Peter Arlet1,2, Jesper Kjær3, Karl Broich4 and Eamer Cooke5

We outline our vision that by 2025 the use of real-world evidence will have been enabled and the value will have been established across the spectrum of regulatory use cases. We are working to deliver this vision through collaboration where we leverage the best that different stakeholders can bring. This vision will support the development and use of better medicines for patients.

Real-world data (RWD) and real-world evidence (RWE) are already used in the regulation of the development, authorisation, and supervision of medicines in the European Union. Their place in safety monitoring and disease epidemiology are well-established while their evidentiary value for additional use cases, notably for demonstrating efficacy, requires further evaluation.1 During the coronavirus disease 2019 (COVID-19) pandemic, RWE rapidly provided invaluable evidence on drug safety, vaccine safety, and effectiveness and we were reminded of the importance of robust study methods and transparency.2 Our vision, anchored in the European Medicines Regulatory Network (EMRN) strategy to 2025, is that by 2025 the use of RWE will have been established and the value will have been established across the spectrum of regulatory use cases.3 Delivering this vision will support the development and use of better medicines for patients.

In December 2018, the US Food and Drug Administration (FDA) published its framework for RWE underpinned by three pillars: whether RWE is fit for use, whether the study design can provide adequate evidence, and whether the study context meets regulatory requirements.4 In 2019 to the European Union, we published the OPTIMAL Framework for RWE, also consisting of three pillars: operational, technical, and methodological.5 More recently, the EU approach places RWE in the wider context of big data and is guided by the priority recommendations of the Big Data Task Force. These recommendations are being implemented through the Big Data Steering Group and the second multi-annual work plan was published in August 2021. Figure 1 represents the workplan with its 11 workstreams which will deliver our vision for RWE by 2025. The workplan places emphasis on collaboration across stakeholders and with international regulatory partners. This work also needs to be seen in the wider EU policy context, most notably the European Commission’s plans for a European Health Data Space.6

Acknowledging different frameworks to conceptualise the challenges and opportunities of RWE, we believe the two main priorities for the European Union are to enable its use and establish its value for regulatory decision making. The EMRN is working to deliver on both priorities through a collaborative approach where we leverage the best that different stakeholders can bring, and where those stakeholders can complement the central role of industry in generating evidence.

ENABLING USE

To enable use, we are working on multiple fronts with our stakeholders, including patients, healthcare professionals, industry, regulatory and public health agencies, health technology assessment bodies, payers, and academia. We are initiating work to establish a data quality framework, not just for RWD but for all data used in regulatory decision making. We are striving to improve the discoverability (findability) of RWD through agreement of metadata for RWD and through a public catalogue of RWD sources that can be used. We are also working on a more robust and comprehensive regulatory framework for RWE.

In the recent past, we have seen an increasing focus on the role of RWD in supporting health technology assessment. The European Medicines Agency (EMA) and some national health authorities have begun to make use of RWD in their decision-making processes, which has led to increased collaboration between industry and regulators. The EMA and its Member States have been actively involved in the development and implementation of methods to support the use of RWD in regulatory decision making. These methods include statistical methods for estimating treatment effects and risk assessments.

The European Medicines Agency (EMA) and some national health authorities have begun to make use of RWD in their decision-making processes, which has led to increased collaboration between industry and regulators. The EMA and its Member States have been actively involved in the development and implementation of methods to support the use of RWD in regulatory decision making. These methods include statistical methods for estimating treatment effects and risk assessments.

Figure 1. Big Data Steering Group workplan to 2023. Eleven workstreams to progress the real-world evidence (RWE) vision.
Our vision is that by 2025 the use of RWE will have been enabled and its value will have been established across the spectrum of regulatory use cases. We are committed to working with stakeholders to deliver this vision and in turn to support the development and use of better medicines for patients.
Ensuring the safe and effective use of medical products is not just a European regulatory responsibility...

...it’s a global responsibility for all stakeholders to support
Ensuring the safe and effective use of medical products is not just a European regulatory responsibility...

...it’s a global responsibility for all stakeholders to support

Ensuring the appropriate use of real-world evidence to inform regulatory decision-making is not just a European regulatory responsibility...

...it’s a global responsibility for all stakeholders to support
OHDSI community
We’re all in this journey together...

OHDSI Collaborators
- 2,367 collaborators
- 74 countries
- 21 time zones
- 6 continents

OHDSI Data Network
- 331 data sources
  - 284 EHRs
  - 28 administrative claims
- 34 countries
- 810 million unique patient records

Open community data standards (OMOP CDM)
Methodological research
Open source development (OHDSI tools)
Clinical evidence generation
Current status quo in observational research makes it challenging to build trust in evidence.
Current status quo in observational research makes it challenging to build trust in evidence.
Current status quo in observational research makes it challenging to build trust in evidence

- Curate data
- Select cohorts
- Implement analysis
Current status quo in observational research makes it challenging to build trust in evidence
Current status quo in observational research makes it challenging to build trust in evidence.
Current status quo in observational research makes it challenging to build trust in evidence

- Curate data
- Select cohorts
- Implement analysis
- Disseminate evidence

- Data quality?
- Measurement error?
- Methods bias?
- Methodological concerns

Is the evidence reliable?
Current status quo in observational research makes it challenging to build trust in evidence.
Current status quo in observational research makes it challenging to build trust in evidence

Does the study provide an unbiased effect estimate?
Are the findings generalizable to the population of interest?

Curate data
Select cohorts
Implement analysis
Disseminate evidence

data quality?
measurement error?
methods bias?
programming correct?
logic correct?
ETL correct?

methodological concerns
technical concerns

Is the evidence reliable?
Current status quo in observational research makes it challenging to build trust in evidence.

Does the study provide an unbiased effect estimate?
Are the findings generalizable to the population of interest?

Curate data → Select cohorts → Implement analysis → Disseminate evidence

- data quality?
- measurement error?
- methods bias?
- programming correct?
- logic correct?
- ETL correct?

Is the evidence reliable?

Can the study be fully reproduced?
Does the analysis actually do what the protocol said it would do?

- methodological concerns
- technical concerns
Observational research across data networks increases complexity and raises new questions

Do the results show a consistent effect across the network?
How does heterogeneity across network (in population composition, data capture process, effect estimates) impact interpretation?

Curate data network
Select cohorts within the network
Implement analysis across the network
Disseminate evidence from the network

Data quality?
Measurement error?
Methods bias?
Methodological concerns

ETL correct?
Logic correct?
Programming correct?
Technical concerns

Can the study be fully reproduced across the network?
Is the evidence reliable?

Can the study be fully reproduced across the network?
<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatable</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Reproducible</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Generalizable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Different</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Robust</td>
<td>Identical</td>
<td>Same or different</td>
<td>Same or different</td>
<td>Different</td>
<td>Similar</td>
</tr>
<tr>
<td>Calibrated (controls)</td>
<td>Similar</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Statistically consistent</td>
</tr>
</tbody>
</table>
### Desired attributes for reliable evidence

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatable</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Reproducible</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Generalizable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Different</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Robust</td>
<td>Identical</td>
<td>Same or different</td>
<td>Same or different</td>
<td>Different</td>
<td>Similar</td>
</tr>
<tr>
<td>Calibrated</td>
<td>Similar (controls)</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Statistically consistent</td>
</tr>
</tbody>
</table>

A system for real-world evidence generation based on consistent application of standardized analytics across a standardized data network can be empirically demonstrated to be reliable.
Common data model can enable standardized analytics across a distributed data network.

Source 1 raw data
- Electronic health records

Source 2 raw data
- Administrative claims

Source 3 raw data
- Clinical data

Transformation to common data model

Source 1 CDM

Source 2 CDM

Source 3 CDM

Open-source analysis code

Open evidence
Common data model can enable standardized analytics across a distributed data network

Need confidence in the quality of the data and its transformation

Transformation to common data model

Source 1 CDM

Need confidence in the quality of the analytics tools and their output

Open-source analysis code

Open evidence

Source 1 raw data
- Electronic health records

Source 2 raw data
- Administrative claims

Source 3 raw data
- Clinical data

Source 1 CDM

Source 2 CDM

Source 3 CDM

Need confidence in the quality of the data and its transformation
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

System characteristics:
• Standardized procedures with defined inputs and outputs
• Analysis packages implementing scientific best practices consistently applied across all data partners, generating consistent output for network synthesis
• Reproducible outputs generated by open-source analysis libraries developed and validated with verifiable unit-test coverage
• Pre-specified and objective decision thresholds for go/no go criteria
• Measurable operating characteristics of system performance
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Data quality evaluation
- Research question
- Database diagnostics
  - Pass
  - Fail
  - STOP

Phenotype development and evaluation
- Cohort definitions
- Cohort diagnostics
  - Fail
  - STOP

Analysis reliability evaluation
- Analysis design choices
- Study diagnostics
  - Pass

Network coordination

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Research question

Database diagnostics
- Pass
- Fail
- STOP

Cohort definitions
- Cohort diagnostics
  - Fail
  - STOP

Analysis design choices
- Study diagnostics
  - Pass

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Network Data Quality Issues by CRIM Table
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Data quality evaluation

Research question → Database diagnostics

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Analysis reliability evaluation

Distributed data network, standardized to common data model

Network coordination
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

‘System’ required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Data quality evaluation

Research question

Database diagnostics

Pass

Fail

Phenotype development and evaluation

Cohort definitions

Cohort diagnostics

Pass

Fail

Analysis reliability evaluation

Analysis design choices

Study diagnostics

Pass

Fail

STOP

System required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

- System required elements:
  - Required phenotypes
  - Analysis specifications
  - Decision thresholds

Data quality evaluation
- Research question
- Database diagnostics
  - Fail
  - Pass
- STOP

Phenotype development and evaluation
- Cohort definitions
- Cohort diagnostics
  - Fail
  - Pass
  - STOP

Analysis reliability evaluation
- Analysis design choices
- Study diagnostics
  - Fail
  - Pass
  - STOP

Distributed data network, standardized to common data model

Network coordination

Interface for exploration

Note: The analysis of distributed system components (SI, samples, S1, samples, S2, samples, S3, samples) was conducted with appropriate control procedures.

Key: S1, S2, S3, S4, Si, Sj, S3, S4, S5, S6, S7, S8

Legend: S1, S2, S3, S4, Si, Sj, S3, S4, S5, S6, S7, S8

Source: Data quality evaluation and cohort diagnostics

Analysis design choices

Study diagnostics

STOP
Complementary types of evidence to generate from real-world data

- **Clinical characterization:** What happened to them?
  - observation

- **Patient-level prediction:** What will happen to me?
  - inference

- **Population-level effect estimation:** What are the causal effects?
  - causal inference
Three potential use cases for the support to committees’ decision-making

From a regulatory perspective, RWE aims to support committees’ decision-making in three main areas

<table>
<thead>
<tr>
<th>Use case category</th>
<th>Use case objective</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support the planning &amp; validity of applicant studies</td>
<td>Design and feasibility of planned studies</td>
<td>Representativeness and validity of Completed studies</td>
</tr>
<tr>
<td>Understand clinical context</td>
<td>Disease epidemiology</td>
<td>Clinical management &amp; drug utilisation</td>
</tr>
<tr>
<td>Investigate associations and impact</td>
<td>Effectiveness and safety studies</td>
<td>Impact of regulatory actions</td>
</tr>
</tbody>
</table>

with permission from Peter Arlett

Classified as public by the European Medicines Agency
Mapping regulatory use cases to evidence types

- **Support the planning & validity of applicant studies**
  - Design and feasibility of planned studies
  - Representativeness and validity of completed studies

- **Understand clinical context**
  - Disease epidemiology
  - Clinical management & drug utilisation

- **Investigate associations and impact**
  - Effectiveness and safety studies
  - Impact of regulatory actions

- **Clinical characterization:**
  - What happened to them?

- **Population-level effect estimation:**
  - What are the causal effects?

- **Patient-level prediction:**
  - What will happen to me?
Mapping regulatory use cases to evidence types

Support the planning & validity of applicant studies

- Design and feasibility of planned studies
- Representativeness and validity of completed studies

Understand clinical context

- Disease epidemiology
- Clinical management & drug utilisation

Investigate associations and impact

- Effectiveness and safety studies
- Impact of regulatory actions

Questions that can be informed with real world evidence:

**Clinical characterization:**
What happened to them?

Who are the patients with disease eligible for treatment?
Who are the patients exposed to those treatments?
How often do outcomes occur amongst those patients?

**Population-level effect estimation:**
What are the causal effects?

Is the outcome causally related to exposure to treatment?
How does the risk compare with alternative treatments?

**Patient-level prediction:**
What will happen to me?

Which risks can be actionably predicted with available data?
Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

- **Design and feasibility of planned studies**
- **Representativeness and validity of Completed studies**
- **Disease epidemiology**
- **Clinical management & drug utilisation**
- **Effectiveness and safety studies**
- **Impact of regulatory actions**

**Questions that can be informed with real world evidence:**

- **Clinical characterization:** What happened to them?
  - Who are the patients with disease eligible for treatment?
  - Who are the patients exposed to those treatments?
  - How often do outcomes occur amongst those patients?

- **Population-level effect estimation:** What are the causal effects?
  - Is the outcome causally related to exposure to treatment?
  - How does the risk compare with alternative treatments?

- **Patient-level prediction:** What will happen to me?
  - Which risks can be actionably predicted with available data?
  - Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

Support the planning & validity of applicant studies
- Design and feasibility of planned studies
- Representativeness and validity of Completed studies

Understand disease epidemiology
- Disease epidemiology
- Clinical management & drug utilisation

Investigate effectiveness and safety studies
- Effectiveness and safety studies
- Impact of regulatory actions

Questions that can be informed with real world evidence:

**Clinical characterization:** What happened to them?
- Who are the patients with disease eligible for treatment?
- Who are the patients exposed to those treatments?
- How often do outcomes occur amongst those patients?

**Population-level effect estimation:** What are the causal effects?
- Is the outcome causally related to exposure to treatment?
- How does the risk compare with alternative treatments?

**Patient-level prediction:** What will happen to me?
- Which risks can be actionably predicted with available data?
- Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

Support the planning & feasibility of applicant studies
- Design and feasibility of planned studies
- Representativeness and validity of Completed studies

Understand clinical context
- Disease epidemiology
- Clinical management & drug utilisation

Investigate associations and impact
- Effectiveness and safety studies
- Impact of regulatory actions

Clinical characterization:
- What happened to them?

Population-level effect estimation:
- What are the causal effects?

Patient-level prediction:
- What will happen to me?

Questions that can be informed with real world evidence:
- Who are the patients with disease eligible for treatment?
- Who are the patients exposed to those treatments?
- How often do outcomes occur amongst those patients?
- Is the outcome causally related to exposure to treatment?
- How does the risk compare with alternative treatments?
- Which risks can be actionably predicted with available data?
- Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

Support the planning & validity of applicant studies
- Design and feasibility of planned studies
- Representativeness and validity of Completed studies

Understand clinical context
- Disease epidemiology
- Clinical management & drug utilisation

Investigate associations and impact
- Effectiveness and safety studies
- Impact of regulatory actions

Clinical characterization:
- What happened to them?

Population-level effect estimation:
- What are the causal effects?

Patient-level prediction:
- What will happen to me?

Questions that can be informed with real world evidence:
- Who are the patients with disease eligible for treatment?
- Who are the patients exposed to those treatments?
- How often do outcomes occur amongst those patients?
- Is the outcome causally related to exposure to treatment?
- How does the risk compare with alternative treatments?
- Which risks can be actionably predicted with available data?
- Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

Questions that can be informed with real world evidence:

- Who are the patients with disease eligible for treatment?
- Who are the patients exposed to those treatments?
- How often do outcomes occur amongst those patients?

Is the outcome causally related to exposure to treatment?
- How does the risk compare with alternative treatments?

Which risks can be actionably predicted with available data?
- Which patients are at highest risk of adverse events?
Mapping regulatory use cases to evidence types

**Support the planning & validity of applicant studies**
- Design and feasibility of planned studies
- Representativeness and validity of completed studies

**Understand clinical context**
- Disease epidemiology
- Clinical management & drug utilisation

**Investigate associations and impact**
- Effectiveness and safety studies
- Impact of regulatory actions

**Clinical characterization:**
- What happened to them?

**Population-level effect estimation:**
- What are the causal effects?

**Patient-level prediction:**
- What will happen to me?

Questions that can be informed with real world evidence:
- Who are the patients with disease eligible for treatment?
- Who are the patients exposed to those treatments?
- How often do outcomes occur amongst those patients?

- Is the outcome causally related to exposure to treatment?
- How does the risk compare with alternative treatments?

- Which risks can be actionably predicted with available data?
- Which patients are at highest risk of adverse events?
Level of proactivity in delivering real-world evidence

Reactive

Bespoke

Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs
Level of proactivity in delivering real-world evidence

**Enabled**

Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model

**Reactive Bespoke**

Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs
Level of proactivity in delivering real-world evidence

- **Responsive**
  Enable fast evidence generation by using interface that allow qualified users to set defined input parameters, execute standardized analyses, and view results upon request.

- **Enabled**
  Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model.

- **Reactive Bespoke**
  Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs.
Level of proactivity in delivering real-world evidence

**Prepared**
Produce pre-computed evidence to enable answer retrieval in ‘real time’ by qualified users when requested; standardized analysis packages executed across network generate results ‘at-scale’ across many target, outcome cohorts.

**Responsive**
Enable fast evidence generation by using interface that allow qualified users to set defined input parameters, execute standardized analyses, and view results upon request.

**Enabled**
Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model.

**Reactive Bespoke**
Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs.
Level of proactivity in delivering real-world evidence

**Anticipatory**
Generate and deliver insights without being asked; answer questions before requested by ‘pushing’ relevant pre-computed evidence to potential evidence consumers

**Prepared**
Produce pre-computed evidence to enable answer retrieval in ‘real time’ by qualified users when requested; standardized analysis packages executed across network generate results ‘at-scale’ across many target, outcome cohorts

**Responsive**
Enable fast evidence generation by using interface that allow qualified users to set defined input parameters, execute standardized analyses, and view results upon request.

**Enabled**
Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model

**Reactive Bespoke**
Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs
Level of proactivity in delivering real-world evidence

- **Anticipatory**: Generate and deliver insights without being asked; answer questions before requested by ‘pushing’ relevant pre-computed evidence to potential evidence consumers.

- **Prepared**: Produce pre-computed evidence to enable answer retrieval in ‘real time’ by qualified users when requested; standardized analysis packages executed across network generate results ‘at-scale’ across many target, outcome cohorts.

- **Responsive**: Enable fast evidence generation by using interface that allow qualified users to set defined input parameters, execute standardized analyses, and view results upon request.

- **Enabled**: Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model.

- **Reactive Bespoke**: Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs.

**Standardized dissemination**
+ **Standardized analysis configurations**
+ **Standardized analysis tools**
+ **Standardized data, network execution**
Level of proactivity in delivering real-world evidence

<table>
<thead>
<tr>
<th>Time-to-evidence</th>
<th>Proactivity Type</th>
<th>Description</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~seconds</td>
<td>Anticipatory</td>
<td>Generate and deliver insights without being asked; answer questions before requested by ‘pushing’ relevant pre-computed evidence to potential evidence consumers</td>
<td>Standardized dissemination +</td>
</tr>
<tr>
<td>~minutes</td>
<td>Prepared</td>
<td>Produce pre-computed evidence to enable answer retrieval in ‘real time’ by qualified users when requested; standardized analysis packages executed across network generate results ‘at-scale’ across many target, outcome cohorts</td>
<td>Standardized analysis configurations +</td>
</tr>
<tr>
<td>~hours</td>
<td>Responsive</td>
<td>Enable fast evidence generation by using interface that allow qualified users to set defined input parameters, execute standardized analyses, and view results upon request.</td>
<td>Standardized analysis tools +</td>
</tr>
<tr>
<td>~days</td>
<td>Enabled</td>
<td>Design and execute standardized analysis packages that apply validated statistical libraries with defined input parameters and fixed output to compile summary results across a network standardized to a common data model</td>
<td>Standardized data, network execution</td>
</tr>
<tr>
<td>~weeks, months, years</td>
<td>Reactive Bespoke</td>
<td>Service bespoke project requests by convening team to align on problem statement, author protocol/analysis plan documents, implement statistical programming code to custom specification, execute analysis across databases, iteratively review results and request post hoc analyses, write summary of results as report, and deliver to decision-maker to ensure it meets their needs</td>
<td></td>
</tr>
</tbody>
</table>
A 5-year vision for expanding the proactive use of real-world evidence across regulatory use cases

Support the planning & validity of applicant studies

~seconds
Anticipatory

~minutes
Prepared

~hours
Responsive

~days
Enabled

~weeks, months, years
Reactive Bespoke

2026
2026
2026

today
A 5-year vision for expanding the proactive use of real-world evidence across regulatory use cases

Support the planning & validity of applicant studies

Understand clinical context
A 5-year vision for expanding the proactive use of real-world evidence across regulatory use cases

- Support the planning & validity of applicant studies
- Understand clinical context
- Investigate associations and impact

- Anticipatory
- Prepared
- Responsive
- Enabled
- Reactive Bespoke

~seconds
~minutes
~hours
~days
~weeks, months, years
Expanding the proactive use of real-world evidence for study planning and validity

- ~seconds: Anticipatory
- ~minutes: Prepared
- ~hours: Responsive
- ~days: Enabled
- ~weeks, months, years: Reactive Bespoke

Support the planning & validity of applicant studies

Review

Contemporary use of real-world data for clinical trial conduct in the United States: a scoping review

James R. Rogers,1 Junghwan Lee,1 Ziheng Zhou,2 Ying Kuen Cheung,3 George Hripcsak,1,4 and Chunhua Weng1

Support the planning & validity of applicant studies
Expanding the proactive use of real-world evidence for study planning and validity

Support the planning & validity of applicant studies

- seconds: Anticipatory
- minutes: Prepared
- hours: Responsive
- days: Enabled
- weeks, months, years: Reactive Bespoke

~seconds, ~minutes, ~hours, ~days, ~weeks, months, years

Support the planning & validity of applicant studies

2026

Clinical comparison between trial participants and potentially eligible patients using electronic health record data: A generalizability assessment method

James R. Rogers a, George Hripacak a,b, Ying Kuen Cheung c, Chunhua Weng a,b

Fig. 1. Overview of study methodology.
Expanding the proactive use of real-world evidence for understanding clinical context

- **Anticipatory**
  - ~seconds
  - 2026

- **Prepared**
  - ~minutes
  - 2026

- **Responsive**
  - ~hours
  - 2026

- **Enabled**
  - ~days
  - 2026

- **Reactive Bespoke**
  - ~weeks, months, years
  - 2026

Understand clinical context

---

https://data.ohdsi.org/Covid19CharacterizationCharybdis/
Expanding the proactive use of real-world evidence for understanding clinical context

- Anticipatory
- Prepared
- Responsive
- Enabled
- Reactive
- Bespoke

Understand clinical context

~seconds

~minutes

~hours

~days

~weeks, months, years

2026

2026

2026

2026

2026
Expanding the proactive use of real-world evidence to investigate associations and impact

THE LANCET
Rheumatology

Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study

Jenny C Land, James Zewert, Kristin Kastiar, Talia Durante-Salles, Mano Teneza F Abarb, Feida Alkaisi, Ossid Alar

Articles

Investigate associations and impact

~seconds
Anticipatory

~minutes
Prepared

~hours
Responsive

~days
Enabled

~weeks, months, years
Reactive Bespoke

2026

2026

2026

2026

2026
Expanding the proactive use of real-world evidence to investigate associations and impact

- ~seconds: Anticipatory
- ~minutes: Prepared
- ~hours: Responsive
- ~days: Enabled
- ~weeks, months, years: Reactive Bespoke

Investigate associations and impact

**Original article**

Risk of depression, suicide and psychosis with hydroxychloroquine treatment for rheumatoid arthritis: a multinational network cohort study

Jennifer C. E. Lane, James Weaver, Kristin Kostka, Talita Duarte-Salles, Maria, Thamir M. Als, Juan M. Band, Jill Hardin, Benjamin Sko, Kristine E. Lynn, Henry Morgan, Fredrik Nyberg, Albert Prats-Lleonart, Anthony G. Scott, Marc A. Sucheston, Junqing Xie, Patrick Ryan, consortium
Expanding the proactive use of real-world evidence to investigate associations and impact

Investigate associations and impact

- seconds
  - Anticipatory
- minutes
  - Prepared
- hours
  - Responsive
- days
  - Enabled
- weeks, months, years
  - Reactive Bespoke

Today's view is 2026
A 5-year vision for expanding the proactive use of real-world evidence across regulatory use cases

- Support the planning & validity of applicant studies
- Understand clinical context
- Investigate associations and impact

- Anticipatory
- Prepared
- Responsive
- Enabled
- Reactive Bespoke

- ~seconds
- ~minutes
- ~hours
- ~days
- ~weeks, months, years

Today → 2026 → 2026 → 2026
Concluding thoughts

• Enabling use and establishing value of real-world evidence is a reasonable vision, which requires building trust across evidence generators and consumers
• People and processes need to be augmented with science, technology and engineering
• Community efforts today can enable a more proactive future tomorrow
  – Data network standardization and quality assessment
  – Design of standardized outputs for regulatory use cases
  – Standardized analytic tool development
  – Phenotype development and evaluation
• Open science systems that promote transparency and reproducibility can increase reliability and efficiency
• Regulatory use cases largely involve characterization analyses, have been demonstrated to be feasible, and are ready-to-scale