

White Paper

Real World Database Studies: Eight Key Steps to Success

HARVEEN UBHI, Epidemiologist, IQVIA Real World Solutions

PIA HORVAT, Senior Epidemiologist, IQVIA Real World Solutions

PAOLA NASUTI, Associate Principal, IQVIA Real World Solutions



Table of contents

	Introduction	3
	Current use of real world data (RWD)	3
	Considerations for a retrospective database study	4
	Key steps to a robust retrospective database study	5
	Conclusion	7
	References	8
	About the authors	10



Introduction

This white paper is designed for individuals within the pharmaceutical industry who are looking to use secondary real world data (RWD) sourced from preexisting administrative or clinical databases to answer specific research and business questions. It is intended as a top-level guide on planning, executing and reporting a high-quality retrospective database study to ensure robust results for submission to healthcare decision-makers.

Current use of Real World Data (RWD)

Healthcare decision-makers have become increasingly aware of the need to understand the impact of interventions in the real world setting, using RWD. The greater reliance on RWD in part stems from a gradual shift in the timing of drug appraisals, which are now being conducted much sooner in the product lifecycle. It also reflects growth in models of earlier access to treatment, adaptive licensing and management entry agreements, which generally involve provisional approval contingent on further data. Re-assessment of these treatments using data collected through product use can help to support subsequent decisions. RWD also now plays a key role in informing decisions about appropriate access and reimbursement, better outcomes measurement and drug development decisions across the lifecycle.

Retrospective database studies are of growing interest to decision-makers seeking to measure the effects of treatments in real world use. There are numerous challenges to conducting these studies. However, by applying rigorous epidemiological principles, involving subject matter experts, and using a robust delivery model, companies can significantly enhance the validity of the results. Best practices for conducting retrospective pharmacoepidemiologic studies using

A **database study** is based on RWD sourced from large, electronically accessible databases of pre-existing records, in most cases collected for the routine management of patients in the form of electronic medical records (EMR) or gathered for administrative purposes. However, these databases can also include pharmacy or population-based registries.¹

Real world healthcare databases generally

- Cover a large number of patient lives
- Are created for administrative purposes and not necessarily to address a particular research question
- Have restrictions on access to data, often requiring a third party

RWD have been published in recent years, providing guidance and recommendations on designing, analyzing and documenting the results.²⁻¹² In this white paper, we provide an overview of the main points that companies should consider when initiating a retrospective study to ensure robust, reliable results.

Considerations for a retrospective database study

Given the growing role of real world retrospective database studies and the potential implications of the results, it is important to conduct them with high scientific rigor. We suggest a structured, scientific approach to overcome operational barriers and increase study validity, recognizing the need for a strong epidemiologic design and proven analytical methods to reduce the potential for bias.

It is important to conduct database studies with high scientific rigor. We suggest a structured, scientific approach to overcome operational barriers and increase study validity, recognizing the need for a strong epidemiologic design and proven analytical methods to reduce the potential for bias.

Standard operating procedures (SOPs) are essential to ensure consistent practices in RWD study execution across different teams. They are also key to confirming that all research tasks are conducted in accordance with institutional, local and federal guidance. SOPs are typically linked to working documents that provide step-by-step instructions to project teams on the various study-related activities (protocol development, data management, statistical analysis plan (SAP) development, quality control plans, reporting, etc).

The extent to which RWD can provide valid and useful evidence depends on how the data is extracted, processed and managed and on the appropriate use of expertise in reporting the study findings. Involving a multi-disciplinary team of data science experts

coordinated by an experienced project manager at all stages from design to delivery is key to achieving high-quality research. Each member of the team plays a crucial role in ensuring the smooth running of the study in accordance with a robust delivery model. Typically, this team should consist of

- **Principal investigator (PI)** responsible for conducting the study
- **Database experts/country-specific advisors** with a deep understanding of the data's strengths and weaknesses and knowledge of the local healthcare system
- **Epidemiologists** with expertise in designing retrospective database studies
- **Physicians or medical experts** who can advise on cohort definition, inclusion/exclusion criteria and diagnosis/procedure codes
- **Statisticians and programmers** with expertise in analyzing large and complex healthcare datasets
- **Data managers** responsible for database extraction and management
- **Medical writers** who can report on the study findings in clear, standard language
- **Project manager** responsible for overseeing all stages of the project and for developing and maintaining the project management plan including resources, risk management and quality control

Key steps to a robust retrospective database study

We have identified eight key steps for designing, executing and reporting a retrospective database study. These are outlined below.

1. Define the study objectives. The first important step is to understand the rationale for conducting the study and why the research is being undertaken. The study objectives should be clearly defined and documented in scientific prose as a hypothesis that can be tested or proven, with a description of how this might be achieved through an epidemiological study.

Retrospective Studies: 8 Key Steps



As an example, in the case of a comparative study the objectives should include a clear indication of

- Exposure groups to be compared
- Outcomes in terms of endpoint being used (e.g., hospitalization, death)
- Duration of follow-up
- Any sub-groups of interest

It is important to bear in mind that if the detail contained in the objectives is insufficient, the detail contained in the methodology will be equally insufficient, leaving the project vulnerable in terms of the overall strategy.

2. Identify the data source(s). One of the key criteria for designing a real world study is to ensure that the chosen database contains the necessary data elements to meet the objectives. This assessment calls for both knowledge of the database and a clear vision of the study design, data components and operational requirements. Before selecting a database, the team should give careful consideration to the diverse and heterogeneous nature of RWD and uniqueness in both the content (e.g., parameters available, diagnostic coding) and context of the source (e.g., healthcare setting, purpose of data collection, geographic representativeness, duration of patient enrollment in the database).

In some cases, if the database of interest does not contain the parameters required to meet the study objectives, it may be necessary to obtain supplemental de-identified data (e.g., data on secondary (hospital) admissions, lab results) from other sources. Linking different data sources under a common identifier (e.g., social security number) may help to fill this gap. However, it is important to understand the strengths and limitations involved in using linked de-identified data. Consulting a database expert who has a deep understanding of the sources can inform this process and help to place the database within the context of the healthcare environment.

3. Develop the study protocol. A protocol is a scientific blueprint for the study design, built on formal epidemiological principles. It should contain detailed strategic and operational instructions for designing a retrospective database study (patient selection, database selection, etc). There are various published guidance documents on good protocol writing (e.g., FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data,⁹ ENCePP Checklist for Study Protocols,¹² ISPE Guidelines for Good Pharmacoepidemiology Practices,¹³ EMA Guidance for the format and content of the protocol of noninterventional post-authorisation safety studies¹⁵).

4. Apply for ethics approval. The ethical requirements for a retrospective database study depend on the governance surrounding the database. They also vary from country to country. Obtaining ethics approval can be a lengthy process, ranging from a few weeks (e.g., CPRD in the UK) to a year (e.g., National Registries in Sweden) according to the database. When applying for approval, it is important to consider any local rules and regulations and to bear in mind that ethics approval is specific to the details in the protocol; any major changes to the protocol following approval will require an additional review by the ethics committee.

5. Build the statistical analysis plan (SAP). The SAP provides operational details on the statistical requirements for executing the contents of the study. The SAP should be written by an experienced statistician who is familiar with the database. For a multi-country study (or where "significantly" different data sources are used within a study), we recommend developing a "global" SAP and then individual country-specific (data sourcespecific) SAPs. In addition, a data preparation plan should be created, providing an unambiguous set of instructions to the programming team to operationalize the study objectives.

6. Extract the data. The data extraction should be executed by a database expert with sufficient knowledge

of the structure, content and context of the database as well as a solid understanding of the healthcare system from which the data is derived.

7. Conduct the analyses. Depending on the scope and complexity of the project, preparing and analyzing the data can take between a few weeks and a few months in most cases. The study project manager should ensure that all team members who are assigned to conduct the analysis have the appropriate level of experience, education and training as well as a deep understanding of the content (e.g., parameters available, diagnostic codes used) and context (e.g., local healthcare setting, purpose of data collection) of the data source.

8. Report the results. The reporting of real world studies requires various skillsets and should be based on feedback from a multi-disciplinary team of data scientists (e.g., epidemiologist, medic, statistician, medical writer). Especially in the case of multi-country, multi-database studies, database experts with country-specific experience should also be involved to ensure correct interpretation of differences across geographies. There are various published documents that offer guidance on clear, transparent report writing, including EMA Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies,¹⁶ the STROBE guidelines,¹⁷ the reporting of studies conducted using observational routinely-collected health data (RECORD) statement,¹⁸ and more recently, the reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORDPE)¹⁹ which extend the STROBE and RECORD statements on pharmacoepidemiological studies.

Each step in a retrospective database study can take considerable time, depending on the project scope, number of databases involved and complexity of the study objectives. It is important to have the oversight of an experienced study **project manager** and a strong **quality control process** involving appropriate individuals at every stage of the work. The significant commitment

required of these individuals should always be factored into the project management plan to ensure realistic expectations and adherence to timelines.

Done well, retrospective database studies can provide unique insights into a wide range of issues, including economic burden of disease, compliance and adherence, healthcare resource use, disease incidence and prevalence, and survival.

Conclusions

When done well, database studies can serve as a supplement to randomized clinical trials (RCTs), providing unique insights into a wide range of issues, including economic burden of disease, compliance and adherence, healthcare resource use, disease incidence and prevalence, and survival. Their planning and execution require time, effort and a strong team of data science experts spanning across multiple disciplines, including epidemiology, statistics, programming, project management and deep knowledge of the healthcare database and the local healthcare setting.

References

1. Using Routinely Collected Data to Inform Pharmaceutical Policies. Analytical Report for OECD and EU countries. https://ec.europa.eu/health/sites/health/files/policies/docs/2019_collecteddata_pharmaceuticalpolicies_report_en.pdf. Accessed on September 27, 2019.
2. Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A checklist for retrospective database studies – report of the ISPOR task force on retrospective databases. *Value Health*. 2003;6(2):90-97.
3. Berger ML, Mamdani M, Atkins D, Johnson ML. Good Research Practices for Comparative Effectiveness Research: Defining, Reporting and Interpreting Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report – Part I. *Value Health*. 2009;12(8):1044-1052.
4. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good Research Practices for Comparative Effectiveness Research: Approaches to Mitigate Bias and Confounding in the Design of Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force – Part II. *Value Health*. 2009;12(8):1053-1061.
5. Johnson ML, Crown W, Martin BC, Dormouth CR, Siebert U. Good Research Practices for Comparative Effectiveness Research: Analytic Methods to Improve Causal Inference from Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report – Part III. *Value Health*. 2009;12(8):1062-1073.
6. Garrison LP, Towse A, Briggs A, de Pouvourville G, Grueger J, Mohr PE, et al. Performance-Based Risk-Sharing Arrangements – Good Practices for Design, Implementation, and Evaluation: Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force. *Value Health*. 2013;16(5):703-719.
7. Berger ML, Martin BC, Husereau D, Worley K, Allen JD, Yang W, et al. A Questionnaire to Assess the Relevance and Credibility of Observational Studies to Inform Healthcare Decision Making: An ISPORAMCP-NPC Good Practice Task Force Report. *Value Health*. 2014;17(2):143-156.
8. Public Policy Committee. International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP). *Pharmacoepidemiol Drug Saf*. 2016;25(1):2-10.
9. US Food and Drug Administration (FDA): Guidance for Industry and FDA Staff: Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. May 2013. <https://www.fda.gov/media/79922/download>. Accessed on September 22, 2019.
10. US Food and Drug Administration (FDA): Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment. March 2005. <https://www.fda.gov/media/71546/download>. Accessed on September 24, 2019.

11. European Medicines Agency (EMA): Guideline on good pharmacovigilance practices (GVP). Module VIII – post-authorisation safety studies (Rev 3). October 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129137.pdf. Accessed on September 22, 2019.
12. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP). ENCEeP Checklist for Study Protocols (Revision 4). http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml. Accessed on September 26, 2019.
13. International Society for Pharmacoepidemiology (ISPE). Guidelines for Good Pharmacoepidemiology Practices (GPP). Revision 3, June 2015. <https://www.pharmacoepi.org/resources/policies/guidelines-08027/>. Accessed on September 26, 2019.
14. STROBE Statement. STrengthening the Reporting of OBServational studies in Epidemiology (STROBE) Checklist. <https://www.strobe-statement.org/index.php?id=strobe-home>. Accessed on September 28, 2019.
15. European Medicines Agency (EMA). Guidance for the format and content of the protocol of noninterventional post-authorisation safety studies. September 2012.
16. European Medicines Agency (EMA). Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies. July 2013.
17. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al, for the STROBE Initiative. STrengthening the Reporting of OBServational studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Med. 2007;4(10): e297. <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0040297>. Accessed on September 27, 2019.
18. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, et al. The Reporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med. 2015;12(10): e1001885. <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001885>. Accessed on September 27, 2019.
19. Langan SM, Schmidt SAJ, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363:k3532. <https://www.bmj.com/content/bmj/363/bmj.k3532.full.pdf>. Accessed on September 27, 2019.

About the authors



HARVEEN UBHI, PHD
Epidemiologist
IQVIA Real World Solutions, UK

Harveen brings several years of experience in designing and conducting retrospective and prospective observational studies. She has published a number of papers in peer-reviewed medical journals and worked on projects across multiple therapy areas including oncology, metabolic disorders, orphan diseases and pain. Prior to joining IQVIA, Harveen was a post doctoral fellow in Epidemiology and Statistics at University College London (UCL). She holds a PhD from Maastricht University and an MSc from Queen Mary University.



PIA HORVAT, PHD
Senior Epidemiologist
IQVIA Real World Solutions, UK

Pia brings several years of experience in designing and conducting retrospective and prospective observational studies. These span ecological, cross-sectional, case-control and multi-centre cohort studies in various chronic disease areas, including cognitive impairment and cardiovascular disease. Pia has published a number of papers in peer-reviewed medical journals and is an experienced data analyst who has used a range of statistical approaches in her work. Prior to joining IQVIA, Pia was a post doctoral fellow in Epidemiology and Statistics at University College London (UCL). She holds a PhD in Epidemiology from UCL and a BA (Hons) in Human Sciences from the University of Oxford.



PAOLA NASUTI, MPHARM
Associate Principal
IQVIA Real World Solutions, UK

Paola joined IQVIA in 2003 after spending a number of years working as a research scientist at Università degli Studi di Camerino (Italy). Over the past 10 years at IQVIA, Paola has managed multi-country retrospective studies using data sourced from claims databases, electronic medical records and registries. She also has years of experience conducting physician surveys. Paola holds a Master's degree in Pharmacy and has completed various postgraduate courses in Epidemiology, Public Health and Medical Statistics (London School of Hygiene and Tropical Medicine, University of Reading). She has co-authored several research articles in peer-reviewed journals and presented at many international medical conferences.



CONTACT US

iqvia.com/contactus

iqvia.com

