Comprehensive Review

PAIN 167 (2026) 8–18

PAIN®

OPEN

Real-world data and evidence in pain research: an IMMPACT comprehensive review and best practice recommendations

Jan Vollert[a,*], John Spivack[b], Blythe Adamson[c,d], Ralf Baron[e], John T. Farrar[f], Ian Gilron[g], David Hohenschurz-Schmidt[h,i], Robert D. Kerns[j], Sean Mackey[k], John D. Markman[l], Michael P. McDermott[m], Michael Parides[b], Andrew S.C. Rice[h], Dennis C. Turk[n], Ajay D. Wasan[o], Robert H. Dworkin[p], Dale J. Langford[q,r], on behalf of ACTTION'S Consortium on Pain Efficacy, Effectiveness, and Safety Studies (COPESS)

Abstract

Real-world data (RWD) can be defined as routinely collected clinical or administrative data that might be used for research purposes and to generate real-world evidence (RWE). Computerized search and data mining methods, large electronic databases, and the development of novel computational and statistical methods allow for improved access to and analysis of RWD. Although RWD afford the opportunity to generate RWE with potentially improved efficiency and generalizability over prospective clinical studies, it is important to understand and apply best practices when analysing RWD, particularly when the goal is to generate RWE of diagnostic, prognostic, or treatment effectiveness. Real-world evidence can provide evidence complementary to randomized clinical trials (RCTs), especially in scenarios where RCTs are difficult to conduct. Real-world evidence studies need to be carefully designed, the research question clearly defined and addressable with the available RWD source, variables (treatment, outcome, covariates) operationalized, and hypotheses and analyses specified before data access. Sound interpretation of results requires a deep understanding of the benefits and limitations of RWE studies, including often deficient data quality, confounding, and other potential sources of bias. Registered protocols, registered reports as a publishing model, and/or restricted access to data until protocols are in place can be encouraged by journals and enforced by data guardians and will contribute to the emergence of high-quality RWD studies. Here, we summarize guidance documents on generating RWE of treatment effectiveness or comparative effectiveness, discuss the strengths and limitations of RWD and RWE, and provide recommendations for conducting effectiveness RWE studies in the pain field.

Keywords: Real-world data, Real-world evidence, Electronic health record, Registry, Registries, Real-world, Historic control

1. Introduction

Real-world data (RWD) are commonly defined as "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources"[4] that were not primarily collected for research (see glossary in **Table 1** for key terms). Repurposing medical data collected as part of routine care to support clinical, policy, and regulatory decision making is a growing interest. Widespread adoption of electronic health record (EHR) systems

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Members of ACTTION'S Consortium on Pain Efficacy, Effectiveness, and Safety Studies (COPESS) are listed in the Acknowledgements section

[a] Department of Clinical and Biomedical Sciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom, [b] Biostatistics and Bioinformatics Program, Hospital for Special Surgery Research Institute, New York, NY, United States, [c] Flatiron Health, New York, NY, United States, [d] The Comparative Health Outcomes, Policy and Economics (CHOICE) Institute, University of Washington, Seattle, WA, United States, [e] Division of Neurological Pain Research and Therapy, Department of Neurology, University Hospital of Schleswig-Holstein, Campus Kiel, Germany, [f] Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA, United States, [g] Departments of Anesthesiology & Perioperative Medicine, Biomedical & Molecular, Sciences, Centre for Neuroscience Studies, and School of Policy Studies, Queen's University and Kingston Health Sciences Centre, Kingston, ON, Canada, [h] Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, United Kingdom, [i] Research Department, University College of Osteopathy, London, United Kingdom, [j] Department of Psychiatry, Yale School of Medicine, New Haven, CT, United States, [k] Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA, United States, Departments of [l] Neurosurgery and, [m] Biostatistics and Computational Biology, University of Rochester, Rochester, NY, United States, [n] University of Washington, WA, United States, [o] Department of Anesthesiology and Perioperative Medicine, University of Pittsburgh, Pittsburgh, PA, United States, [p] Department of Anesthesiology and Perioperative Medicine, Neurology, and Psychiatry, University of Rochester, Rochester, NY, United States, [q] Department of Anesthesiology, Critical Care & Pain Management, Pain Prevention Research Center at Hospital for Special Surgery, New York, NY, United States, [r] Department of Anesthesiology, Weill Cornell Medicine, New York, NY, United States

**Corresponding author. Address: Department of Clinical and Biomedical Sciences, Faculty of Health and Life Sciences, University of Exeter, St Luke's Campus, 79 Heavitree Rd, Exeter EX2 4TH, United Kingdom. E-mail address: j.vollert@exeter.ac.uk (J. Vollert).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

http://dx.doi.org/10.1097/j.pain.0000000000003798

Table 1

Glossary of key terms.

Term	Definition
False discovery rate (FDR)	The expected proportion of incorrect findings compared with all findings claimed (ie, expected proportion of type I errors)
HARKing	Hypothesizing after results are known; formulation of hypotheses after conducting exploratory data analyses
Historical control	Control group composed of patients who received standard of care during some select timeframe preceding availability or assessment of active treatment group; historical controls can replace or supplement a control group
Instrumental variable	Variable used in analyses of observational data to control for (unmeasured) confounders in an effort to construct valid estimates of treatment effects
Propensity score	Probability of patient receiving treatment given a set of baseline characteristics, used to adjust for confounding factors
Real-world data	In the context of healthcare data, data that are collected as part of routine care (eg, electronic health records, administrative databases, patient registries)
Real-world evidence	Evidence (typically related to benefits and potential harms of a treatment) that is derived from analyses of RWD

RWD, real-world data.

and electronic administrative databases (eg, insurance claims data), as well as computational and statistical advances allowing for query and analysis of "big" data, has enabled secondary research use of RWD. With a clearly developed and communicated research question, accessible data source(s), meaningful data, and the necessary skills to query, clean, and analyse RWD, using RWD to generate valid real-world evidence (RWE) of treatment effectiveness and/or comparative effectiveness offer a seemingly efficient approach to advancing pain research.

Pain as a biopsychosocial entity is complex in presentation and can be difficult to manage, and existing treatments are typically only modestly effective and/or have unfavourable adverse effect profiles.[81] Despite the critical need for safe and effective nonopioid analgesics, successful approval of low-abuse potential or novel formulations of analgesics is rare.[56] Although randomized clinical trials (RCTs) are considered the gold standard for determining treatment efficacy, insufficient evidence of efficacy is a common reason for halted development.[45] Aside from being expensive and time- and resource-intensive, RCTs are notoriously challenging regarding recruitment and retention and, in an effort to minimize heterogeneity, impose such narrow inclusion criteria that the study sample rarely adequately reflects the patient population at large or clinical practice. Limited sample size, treatment duration, and loss-to follow-up can also limit the ability to detect sustained benefits and rare or long-term adverse events. In some scenarios, RCTs are difficult to conduct, often in nonpharmacological interventions such as surgery, for example, Although pragmatic clinical trials seek to address some of these challenges, there is growing appreciation for the role of RWD in evaluating treatment effectiveness as a complement to or even partial replacement for RCTs.[32]

Real-world evidence is increasingly recognised as a relevant evidence source to regulators,[4] including the European Medicines Agency[27] and the United States Food & Drug Administration.[82] Correspondingly, the private sector recognizes the value and opportunity of leveraging RWE for decision-making.[19] Understanding the potential utility of RWD in generating RWE of treatment effectiveness, professional societies such as the International Society for Pharmacoeconomics and Outcomes Research and the International Society for Pharmacoepidemiology have published a series of position papers and guidelines to improve transparency in the conduct and reporting of "hypothesis evaluating treatment effectiveness" RWE studies.[9,63,92] However, existing guidance documents often focus on pharmacological questions or surgical procedures, while pain and many related symptoms are likely less clearly and consistently recorded in databases.

To understand current practices for studies using RWD to advance pain research, we previously conducted a review,[84] in which we documented frequent use of RWD, but highlighted the need for improvement in methods and reporting. Pain provides unique challenges for RWD studies: a lack of standard instruments may lead to heterogenous recording, and perhaps more likely, that many—if not most—sources may not contain *any* measure of pain intensity.

Although the use of RWD affords a potentially powerful opportunity to evaluate real-world effectiveness or comparative effectiveness of pain treatments, there are important considerations to address to generate valid and informative RWE.[93] Therefore, the objective of the current review was to provide an overview of existing guidance on the conduct, analysis, and reporting of RWE studies, and to provide recommendations tailored to the field of pain research, with a focus on effectiveness studies.

2. Methods

To identify guidance, best-practice, and methodological documents on RWE, we conducted a PubMed search using the search string:

("Real world" [Title]) AND ((booksdocs [Filter] OR comment [Filter] OR governmentpublication [Filter] OR guideline [Filter] OR introductoryjournalarticle [Filter] OR review [Filter] OR scientificintegrityreview [Filter]) AND (humans [Filter])).

Additional articles were included by searching the reference list of the included articles. Consensus articles, systematic reviews of methods, narrative reviews, and topical reviews were included, while commentaries on single studies were excluded. A total of 41 articles were included.[3,5–7,11,15,17,19,20,22,25,30,31,33,38,39,37,39,46,47,52,50,47–55,52,54,61,58,67,65,73–75,88,90–92,96–98] A detailed description of the search strategy, including PRISMA flowchart, inclusion and exclusion criteria, and included studies are provided in the supplementary material, http://links.lww.com/PAIN/C377.

Considerations and recommendations were formulated based on: (1) our previous scoping review on the use of RWD in pain research studies; (2) the current review of best practices and guidelines for RWE studies; (3) discussion among a diverse multidisciplinary group with expertise in data science, RWE study methodology, clinical trial methodology, and pain research; and (4) input from members of ACTTION's Consortium on Pain Efficacy, Effectiveness, and Safety Studies (COPESS).

3. Search findings

A synopsis of the reviewed studies is provided in the Supplement (Table S1, http://links.lww.com/PAIN/C377). Below, we provide an overview of important considerations—including Definitions, Limitations, and Recommendations—for the conduct of RWE studies in the setting of pain, drawn from the reviewed studies

and the sources described above (ie, previous scoping review, discussion with content experts across multiple disciplines, input from ACTTION's pain consortium).

3.1. Definitions of real-world evidence and real-world data

Although definitions of RWD and RWE can vary,[57] in this article, we use the US Food and Drug Administration's RWD definition and focus on the reuse of existing "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources."[4] However, an alternative use of "real-world" is often to distance what we call prospective "pragmatic trials"—trials conducted in routine practice outside of typical research settings, which still may be randomized or controlled—from RCTs, a field and distinction we have covered before (**Table 2**).[43,44] Ultimately, explanatory RCTs, pragmatic trials, and RWD studies can complement each other[25] and contribute to a constantly refining model of evidence[29,31] and postapproval monitoring.[17,59] Real-world evidence is usually considered as evidence generated from RWD.

3.2. Real-world data for (comparative) effectiveness

Here, we focus on the use of RWD to create RWE of effectiveness. Studying effectiveness of pain treatments is challenging given the inherent variability of pain ratings and observed low-to-moderate effect sizes of treatments that can result in a low signal-to-noise ratio. This affects RCT and RWD studies similarly as well as in specific ways: RCTs cannot increase sample sizes beyond a certain limit to detect smaller signals due to spiraling cost. However, we can only model what we expect to be relevant: Unlike randomization, which is used to reduce known and unknown biases, modelling can only account for known or suspected influential factors.

Real-world data can be used to model comparative trials as described in detail below, or it can be used to enrich explanatory RCTs or pragmatic trials[5,6,28,65]; for example, using usual care data in pragmatic trials.[98]

One of the opportunities of using RWD can be increasing generalizability of findings beyond strictly defined and often self-selected populations of RCTs. Targeted approaches can be used to specifically include underrepresented or minoritized populations (eg, regarding race or ethnicity). However, given that most databases used are currently based in Europe and North America, representation of and generalizability to other populations (eg, African and South(-East)-Asian populations) may be challenging.

3.3. Limitations of real-world data

There are generally accepted limitations of RWD that should not necessarily preclude the use of RWD, but need to be taken into consideration when appraising the value of RWE.[54]

3.3.1. Data integrity

For any given data source, it may be challenging to ensure that data are accurate, reliable, consistent, or complete.[22] Even registries that are prospectively planned and generally ensure higher data quality[20] are not free of the factors that compromise data integrity in nonregistry data (eg, incomplete data, biased entry into registry, nonadjudicated outcomes).[70]

Data are usually not collected in a systematic way, by a limited number of self-selected sites, introducing bias in the population and calling into question the "real-world" external validity that is generally proposed as one of the greatest strengths of RWD.[17] Pain is often incompletely recorded in International Classification of Disease codes, and assessment of pain intensity is often absent. The post hoc nature of data assessment means that study design is guided by the available data rather than vice versa, limiting the questions that can meaningfully be investigated. Without commonly accepted frameworks for rigorous analysis and reporting, data cleaning and analytic methods are not always following best practices; given the natural constraints of using existing, nonrandomized data, this can affect the results. The impossibility to control for all confounding factors can introduce substantial bias to the results.[77] Some of these limitations can be addressed, or their impact controlled, for example, by using sensitivity analyses to assess the impact of missing data and selection bias, and use of inferential statistics to address confounding factors where possible.[71] Other more serious limitations, however, such as insufficient pain coding or lack of pain measures, are fundamental and call into question whether meaningful evidence can be derived from the data.

3.3.2. Patient consent, privacy, and anonymity

Patient consent, privacy, and anonymity can pose a general limitation of RWD, as, unlike in prospective trials, RWD will often include data that were not given with specific consent for research. Real-world evidence will be most powerful if the RWD used is multifaceted; however, this will almost certainly mean linking of various data sources or databases, and interoperability between databases (eg, by using common personalised

Table 2

Characteristics of randomized clinical trials, pragmatic trials, and real-world data studies using existing databases.[54,59]

	Randomized controlled clinical trials	Pragmatic trials	Real-world data studies
Strengths	Best for studying efficacy of an intervention	Best for external validity	Large sample, fast, fewer resources required
Weaknesses	High cost, questionable external validity	Blinding, oversight, unclear internal validity	Limited to available data, variable protocol/ treatment procedures, unclear validity
Population	Specific inclusion	Routine care	Usually routine care*
Data integrity	High	Medium	Variable
Data heterogeneity	Controlled/Low	High	Variable
Cost	High	Medium	Usually low†
Validity	Internal	External	Variable

This is a simplification, and there is significant overlap between these concepts: many pragmatic trials are randomized, and most are controlled, eg, a deep dive into pragmatic trials and the distinction to clinical trials can be found in References 43 and 44.

* In some cases, registries might be based on specific inclusion criteria, or population-based sampling.

† Using already collected data, cost is low. Maintaining prospective registries, however, has significant associated costs.

identifiers) will by nature reduce anonymity.[77] Even for prospective inclusion, for example, in registries, using informed consent, special considerations differing from those in RCTs need to be taken into account.[47] Adherence to ethical and legal privacy frameworks is imperative.

3.3.3. Multiple hypothesis testing

Multiple hypothesis testing, inflation of the probability of a false positive result, and HARKing (Hypothesising After Results are Known)[49] affect all research, but the post hoc nature of working with existing RWD enables quick and low-cost testing of multiple hypotheses in a quick sequence with subsequent selective reporting of "significant" results.[71] We have shown previously that most RWD studies in pain research report significant findings,[84] emphasising a need to control for inflation of the probability of false positive findings that can otherwise destroy confidence in RWE. For clinical trials, pre-registration of studies with associated protocols and predetermined primary outcome measures have proven effective in reducing selective reporting. However, this can only partly be implemented for RWD.[71] Registries, for example, can demand such registered protocols before granting data access, but for hospital or insurance data, there is no guarantee that the protocol was designed after tentative results were already known.[19]

All data sets have inherent biases that cannot be completely overcome. For example, there may be a bias in patients willing to permit the use of their data (self-selection bias); in longitudinal studies, time points at which assessment is obtained are fixed and cannot be optimized to the study question and real time-to-event may differ (immortal time bias), or patients might selectively comply with treatments (compliance bias). As with other aspects, these biases can affect prospective trials as well; however, prospective trials can often address such aspects through careful statistical analysis, while RWD studies will have to work with collected data and acknowledge bias that cannot be mitigated.

4. Recommendations for pain real-world data study conduct

A brief summary of our recommendations with reference for further reading is summarized in **Table 3**.

Table 3

Short form recommendations and related resources for further reading.

Recommendation	Resource
Define a research question and objectives. Only with clarity on the research question, ideally using PICOT, will you be able to assess if your data set is suitable	75,78,86,98
Define the patient populations. Clear inclusion and exclusion criteria must be phrased, while as few restrictions as possible should be made to maximize the use of routinely collected data	47,64,91,98
Operationalize the variables. While pain or surrogate variables will be collected in many data sets, not all will be comparable or realistically capture pain. Continue only if you can operationalize a meaningful pain measure for your data set and question	12,24,37
Use appropriate data sources. The various possible data sources for RWD have distinct (dis) advantages, which need to be taken into account at planning stage, especially regarding generalizability of findings	7,20,35,39,58
Carefully choose analysis methods. As RWD studies are not randomized and data collection usually not prospectively planned, pay special attention to inferential statistical methods that can adjust for confounding factors and deal with unexpected heterogeneity	3,15,33,38,39,52,79,88,90,96,98
Ensure reproducibility and report rigorously. Comprehensive reporting, starting by publishing or time-stamping protocols and considering registered reports, is crucial to demonstrate the veracity of your work	19,26,62,65,68
Consider added value. It may be tempting to perform RWD studies just because of the availability of data, but they still cost resources to you, editors, reviewers, and readers. Only if your work will add to evidence rather than noise should you perform it	17,19,22,36,47,54,61,62,68,71,77,78,97

Resources provide further and in-depth reading on the recommendations.
PICOT, patient population, intervention, comparator, outcome, timeframe; RWD, real-world data.

4.1. Define a research question and objectives

Before the initiation of an RWD study, the specific research question should be clearly defined, ideally using the patient population, intervention, comparator, outcome, timeframe (PICOT) approach (ie, describe the Patient population, Intervention, Comparator, Outcome, and Timeframe).[86] Primary (and secondary) end points should be clearly stated. Applying PICOT and main study objectives will afford a roadmap to investigators for their overall study design, including identification of a fit-for-purpose data source and developing an appropriate analytic approach. Protocols should be developed in advance and may follow standardized templates.[86]

4.2. Define the patient populations

As with RCTs, inclusion and exclusion criteria should be clearly defined. Some criteria may depend on the data source, which relates both to the queried database and the sources feeding into the database. Consistent with recommendations for pragmatic trials, RWE studies should include all patients intended for routine clinical care relevant to the research question, with as few restrictions as possible.[53]

4.3. Operationalize the variables

As in all research questions, defining relevant and quantifiable outcomes and outcome measures is crucial, but the importance of this aspect is heightened in both RWD—where we rely on existing data—and pain research, where pain assessments may vary widely. As opposed to standardized biochemical assays, for example, collection of pain outcomes is likely heterogenous and may not only include different measurement instruments of pain (eg, numeric rating scale, visual analogue scale, verbal rating scale) but may also include patients' ratings of "current," "average," or "worst" with different recall periods (eg, past 24 hours, past week, past month). The timing of pain assessment may also vary regarding exposure to the treatment and may or may not be reported as site or condition-specific. Moreover, pain may be reported by the patient directly or through a clinician or significant other and may not always be documented. Therefore, variables that serve as proxies for pain outcomes may be necessary, and investigators should thoughtfully consider whether such proxy outcomes adequately reflect the outcome of interest.

The challenge does not only pertain to measuring pain but also to other variables relevant to pain. For any pain study, akin to IMMPACT recommendations for the conduct of clinical trials of chronic pain treatments,[24,80] investigators should consider other pertinent outcome domains if available (eg, physical and emotional functioning, patient-perceived improvement or satisfaction, adverse events [AEs], treatment adherence).[12,85] In most RWD settings, pain-focussed patient-reported outcomes will not available, especially in claims databases, but also most EHRs will not provide standardised patient-reported outcomes. Adverse events may not be easily extracted or clearly attributed to the pain treatment in question. For treatment adherence, there may be discrepancy between prescription and consumption, especially for as-needed pain medications. Other relevant aspects, such as repeated healthcare utilization or return to work, will not be stringently coded but can potentially be operationalised to a degree: number of visits can be quantified using claims data and in many healthcare systems, insurance claims document absence of work periods. However, these will not be universal, depend on the country, healthcare system, and potentially insurance provider.

Consideration should also be given to potential confounding variables, covariates that may be used to improve precision of treatment effect, and instrumental variables that may serve to control for unmeasured confounding (discussed below in *Analysis Methods*). Availability of specific diagnosis and treatment codes, measurability of confounds, and using outcome measures that are likely to be recorded (eg, hospitalization, prescription/dispensing data) can improve the robustness of the results.[28]

4.4. Use appropriate data sources

Data in real-world studies can originate from various sources, each with its own advantages and disadvantages. These data sources are not strictly separate, and increasingly, data are combined from multiple sources to overcome the limitations of single sources. A comprehensive review of current data sources is outside the scope of this review, but has been performed by Bellows et al.[7] for pain and Hak et al.[39] for orthopaedic research. None of these data sources are generally preferable to the other: a data set ideally suited to answer one specific question might be incompatible to another—the importance is to identify a fit-for-purpose data source. Below, we provide some considerations for the most common sources of RWD: EHRs, prescription/health insurance claims, and registries.[20] **Table 4** summarizes how treatment, pain (or proxy) outcome, covariates, and instrumental variables (discussed further in the *Analysis Methods* section) might be operationalized by data source as well as strengths and limitations.

4.4.1. Electronic health records

Electronic health records, as an umbrella term, includes single-site settings as well as multisite data warehouses.[77] Single-site data will potentially include in-depth case data that cannot be retrieved from the other sources.[20] In case of missing data, there is the potential of follow-up and retrieval, which is principally not possible for the other sources.[20] Electronic health records therefore have the theoretical advantage of high data integrity and depth, although not guaranteed. Single-site EHR data may reduce generalisability of findings to other patient cohorts, and lack of external data integrity control may reduce confidence.[20] Single-site studies will also be more likely to be affected by underlying bias or confounding factors, such as single surgeon or physician performing a procedure or treatmentand the included population being affected by selection bias through geography. In the form of EHR warehouses, which collate and consolidate data across multiple sites can provide large, multisite data sets with potentially high depth and high generalisability.[62] Access can be through hospital information systems such as Epic[39,74] (eg, Epic Cosmos that combines Epic data from participating institutions) or multiple sites within a large healthcare system (eg, Kaiser Permanente Research Bank). With continued advancements in analytic approaches such as natural language processing, clinical notes can provide an unprecedented amount of rich and nuanced patient-level data.[1,51] Despite the potential of multidimensional, high-depth large-scale patient-level data, investigators must still consider the availability of operational, clearly defined variables, as described above.

4.4.2. Claims data

Claims data, or health insurance records, provide "big" data, across multiple clinical sites, with potentially high patient and case heterogeneity. If retrieved from insurance databases, they can include multilevel data, such as diagnoses, prescriptions, and procedures[20] across inpatient and outpatient settings (eg,

Table 4

Possible operationalization of data sources along with strengths and weaknesses.

Data source	Potential treatment variables	Pain or pain proxy outcome variables	Covariates and instrumental variables	Strengths	Limitations
Electronic health records	Visit type(s); procedure code(s); prescription(s)	Patient-reported pain ratings; routinely collected PROMs; pain-related care utilization; engagement in PT; rescue analgesic use	Sociodemographic characteristics (age, sex, gender, race, ethnicity, living situation, partner status, insurance type); comorbid conditions; concomitant pain treatment(s); provider or prescriber; pre-post insurance change or implementation of new treatment guidelines	Rich patient-level data Patient-reported pain	Potential for missing/ fragmented data, loss to follow-up; typically specific to health system (may not capture outside care); decreased generalizability if single site
Administrative/ claims database	Visit type; procedure code(s); Prescription(s)	Pain-related diagnostic code; # of pain-related procedures, analgesic prescriptions/refills	Basic sociodemographic data (age, sex); comorbid conditions; concomitant pain treatments; treatment adherence; prescriber; pre-post new treatment guideline	Large amount of data; captures data across health systems; typically complete data for all billable encounters	Limited patient-level data; loss to follow-up with insurance changes; decreased generalizability (given all insured)
Registries	Depends on captured data; documented treatment(s)	Depends on captured data; pain ratings, pain-related patient-reported outcome measures	Depends on captured data	Prospectively planned, well-defined data; data quality/ validation procedures likely in place	Registry entry criteria may introduce bias; may be limited predefined data (although may be possible to link with EHR)

EHR, electronic health record, PROM, patient-reported outcome measure, PT, physical therapy.

MarketScan, Truven). However, pain-related outcomes are usually not systematically collected and recorded and only measured indirectly (eg, prescription of pain medication is used as a proxy for pain). International Classification of Disease codes (up to and including version 10) covered few pain diagnoses, many codes and terms are used interchangeably, and pain may not be coded directly. For example, a patient with low back pain due to spinal stenosis may only get a diagnosis code for billing purposes by the spine surgeon at the time of a laminectomy as "spinal stenosis." Adherence is unclear, as is off-label use of prescription medications (eg, antidepressants or anticonvulsants that may be used to treat pain), and anything not claimed or prescribed, such as self-paid medications or treatments. Moreover, by their nature, insurance claims databases do not include uninsured patients, thus introducing selection bias and limiting generalisability to a vulnerable patient group. Nonetheless, the sheer volume of data can provide unique insights, particularly for rare or understudied populations, such as people living with rare pain conditions or vulnerable populations (eg, pediatric patients). Large, national administrative databases, can also be leveraged to conduct population-based studies; for example, the French nationwide healthcare database has been used to provide insights into the prevalence of chronic pain in special populations.[16,23,48]

4.4.3. Registries

Registries are usually prospectively designed to collect and collate data pertaining to general or specific questions,[20] and may include data collected outside of routine care, but can also automatically siphon data from EHRs for registries focused on a particular population of interest. Regardless, registries are typically actively managed and often require specific research proposals to share relevant data with investigators. Registries designed for narrow fields or specific questions will include most data relevant to the respective field, while broader registries can aim to be representative for certain populations.[55] Associated data dictionaries can provide insights into the availability and type of pain-related variables, and as above, the registry must be fit-for-purpose (ie, provide the necessary information to answer relevant research questions). Given the effort and cost to maintain, registries can be limited by funding, and access may be restricted or fee-based. Entry of data into registries that collect data outside of routine clinical care may require informed consent of patients, thereby limiting the patient inclusion.[58]

Two examples of registries with a specific focus on pain are PAIN OUT and CHOIR.

4.4.4. PAIN OUT

PAIN OUT captures detailed pre-, intra-, and acute postoperative clinical data, as well as patient-reported outcomes (based on the International Pain Outcomes questionnaire), with a focus on acute postoperative pain.[95] Since 2019, more than 300 hospitals have contributed data worldwide, with data on more than 500,000 patients currently included. Approximately 100 peer-reviewed publications have been published based on data generated from PAIN OUT, having led to clinically relevant improvements in pain management.[33,50,94]

4.4.5. CHOIR

CHOIR is an example of a clinically embedded, open-source, flexible, registry-based learning health system; developed to provide high-quality, point-of-care data for optimizing care and facilitating real-world research.[13,73] CHOIR has also served as a platform for pragmatic clinical trials involving real-world patients.[21,67] Although initially designed for adult chronic pain clinics, CHOIR has been adapted for other medical specialties, including pediatric pain,[10] interventional radiology,[42]

gastrointestinal medicine,[38] and primary care/family medicine.[40] A notable advantage of a clinically embedded system such as CHOIR is its low data entry cost, as no additional staff are required for data collection. However, a potential limitation is that most data are collected in association with clinical appointments, resulting in temporal asynchrony between data collection points. Two examples of comparative effectiveness studies produced by CHOIR focussed on embedded mental health services vs standard of care[34] and medical cannabis vs prescription drugs.[89] Currently, it is being used to conduct a comparative effectiveness study investigating 3 methods for reduction of opioids as treatment for chronic pain (cognitive behavioural therapy, group self-management, and taper only).[21]

4.5. Carefully choose analysis methods

All populations skew in unique ways—database populations may be especially affected by unexpected distributions because of differing inclusion criteria in multiple sites, eg, it is important to report and analyse data in a way that appropriately captures their heterogeneity instead of overly relying on mean-centered approaches.

Despite recent impressive advances in machine learning and AI approaches, we do not recommend training AIs to be used in effectiveness studies at the moment. Decision making in effectiveness studies needs to be transparent, a feature contradicted by the inherent "black-box" nature of machine learning. Machine learning algorithms have worked exceptionally well in finding hidden patterns, solving complex puzzles, and uncovering the weakest signal from the noise.[4] The goal of effectiveness studies, on the other hand, is to demonstrate a clear and robust signal, for which machine learning may not provide value over rigorous traditional statistical approaches. That being said, recent advancements in machine learning, such as natural language processing, will improve the depth and scope of RWD available for analyses (eg, physician notes), and, with an eye toward personalised medicine, machine learning may be applied to identify patient subgroups or phenotypes associated with treatment effect (including perhaps a more sophisticated multidimensional outcome that encompasses treatment benefit with adverse events and treatment cost).[79]

Toh[79] separates issues pertaining to analysis of RWD studies into 5 domains: study design, exposure type, outcome type, covariate summarization technique, and covariate adjustment method. Study designs include both between-individual designs, such as cohort and case–control designs, as well as intra-individual designs, such as longitudinal and before–after studies. Exposure and outcome types affect the choice of analysis methods, as they can be binary, categorical, ordinal, or continuous.[79] The choice of a statistical model should accurately reflect the exposure and outcome variables.

As interventions in RWD are not randomized, appropriate statistical adjustment for covariates, especially unbalanced confounders of purported treatment effects, is of the highest importance to yield valid RWE. Inclusion of long lists of covariates and complicated functional forms into statistical models can increase the complexity and undermine credibility of such models. If multiple covariates are relevant, propensity score-based methods can be used to accomplish valuable covariate summarization for practical analyses.

4.5.1. Propensity score techniques

Propensity score techniques are a set of statistical methods used to adjust for confounding factors in comparisons between treatment groups.[52] They are particularly useful when adjustment for multiple covariates is necessary and relatively few cases are included. A propensity score is a "balancing score" that can be used to correct for uneven covariate patterns in the treatment groups of an experiment.

Mathematically, the propensity score is derived as the (estimated) probability of membership of a particular intervention group given a set of potentially confounding baseline variables. It reduces the possibly large set of potential confounders (eg, age, social status, ethnicity, sex) to a single variable (or $k - 1$ variables if there are k intervention groups).[18] Decision on covariates to be included in propensity scores can be made based on expert judgment or be data-driven.[52] After constructing the propensity score(s), they can be used in various ways to adjust for potential confounding without having to explicitly include each of these confounders in the statistical model. Methods include propensity score matching, stratification, and inverse probability weighting.[18,52] Loke and Mattishent[52] provide a practical recommendations for applying propensity score-based methods.

Propensity score techniques are key methods for addressing the issue of confounding present in RWD. It should be noted that these methods can only be used to adjust for known covariates, and we should assume that there are unknown, and unmeasured, covariates with the potential to bias the results as well. Sensitivity analyses help determine the potential degree of influence of unmeasured covariates on the results[96] and assessing consequences of violations of assumptions of statistical models[83]; they are critical for appraising the level of evidence that can be generated from an RWD study.

4.5.2. Instrumental variable analyses

Instrumental variable analyses are methods to construct valid estimates and inferences on treatment effects in the presence of confounding factors.[35] Notably, some of these confounding factors may be unmeasured.

Specifically, an instrumental variable Z is an additional variable used to estimate the causal effect of a variable X on an outcome Y. The traditional definition qualifies a variable Z as an instrument (relative to the pair [X, Y]) if (1) Z is independent of all variables (including error terms) that have an influence on Y that is not mediated by X and (2) Z is not independent of X.[66] Therefore, the instrumental variable Z affects Y only through its effect on X.

In the setting of pain research, possible instrumental variables might include prescriber preference for a particular therapy, insurance or change in insurance (to account for coverage of treatments by some payers but not others), or changes in legal framework or implementation of new treatment guidelines (eg, Centers for Disease Control and Prevention opioid prescribing guidelines for chronic pain in 2016). However, in many situations, readily available instrumental variables may not exist due to violations of conditions (1) or (2).[2] The potential of instrumental variables methods to address analysis issues connected with RWD is promising, and perhaps underused. Intriguingly, recent work of the RCT-DUPLICATE initiative has shown its capacity to emulate clinical trials.[30,87]

To account for multiplicity of hypotheses testing in prospectively planned analyses, the most conservative approach is to apply Bonferroni correction, adjusting the alpha level globally by splitting it equally for the number of tests performed. However, a less conservative and preferable method that is the standard practice is to control the family-wise error rate at a prespecified alpha level (eg, 0.05).[72] This more sensitive method controls the chance of at least one

false-positive result and applies to RWD-based studies in the same manner as it does to other studies.[41]

4.6. Ensure reproducibility and report rigorously

To generate RWE, studies using RWD must be reported transparently and comprehensively.[19] Patorno et al.[65] have suggested reporting guidelines for RWE studies that should be considered in addition to the standard reporting guidelines for observational studies, such as The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)[26] guideline and its extension for routinely collected data, RECORD (REporting of studies Conducted using Observational Routinely collected health Data).[8] To evaluate the quality of RWD, O'Leary and Cavender[62] provide criteria in 6 general domains: completeness of data, transparency of data, data audit, data accuracy, patient rights and privacy, and fitness for purpose. It will be a task for journals to set minimum quality requirements for RWE studies to be published.

4.6.1. Registered reports

Registered reports may be a particularly interesting concept for RWE studies. Registered reports are written protocol, which are peer-reviewed like a full article, before the analyses are conducted. Once peer review is conducted, the protocol is published and, at a later stage, results and discussion are added.[14] This concept is meant to reduce publication bias, and while not currently being offered by any specialist pain journal, it has found some uptake in general biomedical journals; for example, *PLoS Biology*, *Cortex*, and multiple *Nature* family journals, such as *Nature*, *Nature Human Behaviour*, and *Nature Neuroscience*.[14,60] Initial results indicate that registered reports have higher methodological quality compared with regular publishing concepts.[76]

4.6.2. Publishing and registering time-stamped protocols

Publishing and registering time-stamped protocols may be more easily implemented to improve reporting and mitigate hypothesising after results are known (HARKing).[63] Protocols can be published on the Open Science Framework (osf.io), where they can be time stamped and either published immediately or remain private until the related manuscript is submitted. Access to EHR warehouses and registries can be restricted until protocols are written and agreed upon to reduce risk of HARKing.

Transparent reporting of RWE studies will not only address HARKing but will also facilitate reproducibility and aggregation of data across studies.[9]

4.7. Consider added value

Ultimately, RWD studies must be conducted to derive meaningful evidence. This goal can only be achieved if the study is adequately designed, and the accessible data are appropriately suited to answer the question. If limitations cannot reasonably be overcome, and therefore meaningful evidence cannot be derived, the study should not be conducted. The fact that RWD studies are low cost does not mean they should be performed regardless of value gained. We fully endorse the recently published white paper on Enhancing Trust in Pain Evidence (ENTRUST-PE),[61] which provides guidance to follow before, during, and after conduct of studies.

Specifically, we would propose that the following questions be answered with a "yes" to conduct an RWD study focussing on effectiveness in pain:

(1) Is there a meaningful pain measure in the data set?
(2) Does the data set extrapolate to wider populations?
(3) Are all relevant known confounding variables captured in the data set and can be adjusted for?
(4) Can inferential statistical analyses be feasibly performed?

5. Conclusions

Leveraging RWD to study treatment effectiveness or comparative effectiveness presents an intriguing opportunity to generate RWE that is broadly generalizable, and that has the ability to study rare pain conditions and detect rare events, and relatively low cost, thereby addressing challenges germane to traditional RCTs.

Given increased availability and anticipated use of RWD to generate RWE, it is critical that investigators apply a responsible approach and thoughtfully (and proactively) consider potential pitfalls and sources of bias that can be present among studies using RWD.[96] Some considerations are particularly relevant to pain given the heterogenous reporting of pain in databases and the biopsychosocial complexity of reporting pain that may not be adequately captured by available data. Investigators should aim to demonstrate that data are of high and uniform quality, and that biases such as self-selection, self-reporting, measurement are not a risk to the validity of the conclusions of the study. It is also incumbent on investigators to demonstrate that best practices have been used at all stages from conception to design to analysis of an RWE study to mitigate such biases. All these challenges will require multidisciplinary teams, including clinical researchers with content expertise, navigating pitfalls in data coding, as well as data specialists and statisticians to ensure rigorous study design.

Despite their limitations, we advocate further efforts to advance responsible use of RWD for the assessment of treatment effects in pain research. The quality of evidence derived from well-executed and well-analyzed RWD-based studies can provide valuable evidence of treatment effectiveness or comparative effectiveness, presenting a valuable opportunity for use by key decision makers, including regulators, to formulate evidence-based recommendations and to influence treatment practices in pain.

Best practices in design, rigorous analysis, and careful attention to potential sources of bias, issues of data quality, and other threats to the validity of conclusions are essential. In the future, to ensure rigorous reporting of RWD studies, a focused extension of the STROBE (The Strengthening the Reporting of Observational Studies in Epidemiology) guideline,[26] clearly setting standards for reporting RWD studies in pain research, would be valuable to further increase the usefulness such work.

Conflict of interest statement

The authors report no conflicts of interest. All interests are reported separately per author.

Acknowledgements

The authors acknowledge ACTTION's Consortium on Pain Efficacy, Effectiveness, and Safety Studies Steering Committee members who reviewed and approved this manuscript, including: Stephen Bruehl, PhD, Robert R. Edwards, PhD, Scott Evans, PhD, Roy L. Freeman, MD, Simon Haroutounian, PhD, Tuhina Neogi, MD, PhD, Kushang V. Patel, PhD, Bryce B. Reeve, PhD, Michael C. Rowbotham, MD, Lee Simon, MD, and Eugene R. Viscusi, MD. Financial support was provided by the Analgesic,

Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) public-private partnership with the United States Food and Drug Administration (FDA), which has received research contracts, grants, or other revenue from the FDA, multiple pharmaceutical and device companies, philanthropy, royalties, and other sources. The views expressed in this article are those of the authors and no official endorsement by the FDA or the pharmaceutical and device companies that provided unrestricted grants to support ACTTION should be inferred.

Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/C377.

Article history:
Received 31 October 2024
Received in revised form 21 June 2025
Accepted 26 June 2025
Available online 3 September 2025

References

[1] Adamson B, Waskom M, Blarre A, Kelly J, Krismer K, Nemeth S, Gippetti J, Ritten J, Harrison K, Ho G, Linzmayer R, Bansal T, Wilkinson S, Amster G, Estola E, Benedum CM, Fidyk E, Estévez M, Shapiro W, Cohen AB. Approach to machine learning for extraction of real-world data variables from electronic health records. Front Pharmacol 2023;14:1180962.
[2] Angrist JD, Krueger AB. Instrumental variables and the search for identification: from supply and demand to natural experiments. J Econ Perspect 2001;15:69–85.
[3] Arora P, Boyne D, Slater JJ, Gupta A, Brenner DR, Druzdzel MJ. Bayesian networks for risk prediction using real-world data: a tool for precision medicine. Value Health 2019;22:439–45.
[4] Baumfeld Andre E, Carrington N, Siami FS, Hiatt JC, McWilliams C, Hiller C, Surinach A, Zamorano A, Pashos CL, Schulz WL. The current landscape and emerging applications for real-world data in diagnostics and clinical decision support and its irpact on regulatory decision making. Clin Pharmacol Therap 2022;112:1172–82.
[5] Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: the changing landscape of the regulatory approval process. Pharmacoepidemiol Drug Saf 2020;29:1201–12.
[6] Beaulieu-Jones BK, Finlayson SG, Yuan W, Altman RB, Kohane IS, Prasad V, Yu KH. Examining the use of real-world evidence in the regulatory process. Clin Pharmacol Ther 2020;107:843–52.
[7] Bellows BK, Kuo KL, Biltaji E, Singhal M, Jiao T, Cheng Y, McAdam-Marx C. Real-world evidence in pain research: a review of data sources. J Pain Palliat Care Pharmacother 2014;28:294–304.
[8] Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement. PLoS Med 2015;12:e1001885.
[9] Berger ML, Sox H, Willke RJ, Brixner DL, Eichler H-G, Goettsch W, Madigan D, Makady A, Schneeweiss S, Tarricone R, Wang SV, Watkins J, Daniel Mullins C. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE special task force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf 2017;26:1033–9.
[10] Bhandari RP, Feinstein AB, Huestis SE, Krane EJ, Dunn AL, Cohen LL, Kao MC, Darnall BD, Mackey SC. Pediatric-Collaborative Health Outcomes Information Registry (Peds-CHOIR): a learning health system to guide pediatric pain research and treatment. PAIN 2016;157:2033–44.
[11] Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. Nat Rev Clin Oncol 2019;16:312–25.
[12] Bova G, Domenichiello A, Letzen JE, Rosenberger DC, Siddons A, Kaiser U, Anicich A, Baron R, Birch J, Bouhassira D, Casey G, Golden K, Iyengar S, Karp BI, Liedgens H, Meissner W, Nicholson K, Pogorzala L, Ryan D, Treede R-D, Tugwell P, Turk DC, Vincent K, Vollert J, Williamson PR, Woller SA, Zaslansky R, Wandner LD, Pogatzki-Zahn EM. Developing consensus on core outcome sets of domains for acute, the transition from acute to chronic, recurrent/episodic, and chronic pain: results of the INTEGRATE-pain Delphi process. eClinicalMedicine 2023;66:102340.
[13] Callahan A, Ashley E, Datta S, Desai P, Ferris TA, Fries JA, Halaas M, Langlotz CP, Mackey S, Posada JD, Pfeffer MA, Shah NH. The Stanford Medicine data science ecosystem for clinical and translational research. JAMIA Open 2023;6:ooad054.
[14] Chambers CD, Tzavella L. The past, present and future of registered reports. Nat Hum Behav 2021;6:29–42.
[15] Chen Z, Liu X, Hogan W, Shenkman E, Bian J. Applications of artificial intelligence in drug development using real-world data. Drug Discov Today 2021;26:1256–64.
[16] Chenaf C, Delorme J, Delage N, Ardid D, Eschalier A, Authier N. Prevalence of chronic pain with or without neuropathic characteristics in France using the capture-recapture method: a population-based study. PAIN 2018;159:2394–402.
[17] Corrao G, Cantarutti A. Building reliable evidence from real-world data: needs, methods, cautiousness and recommendations. Pulm Pharmacol Ther 2018;53:61–7.
[18] D'Agostino RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17:2265–81.
[19] Dagenais S, Russo L, Madsen A, Webster J, Becnel L. Use of real-world evidence to drive drug development strategy and inform clinical trial design. Clin Pharmacol Ther 2022;111:77–89.
[20] Dal Pan GJ. The use of real-world data to assess the impact of safety-related regulatory interventions. Clin Pharmacol Ther 2022;111:98–107.
[21] Darnall BD, Mackey SC, Lorig K, Kao M-C, Mardian A, Stieg R, Porter J, DeBruyne K, Murphy J, Perez L, Okvat H, Tian L, Flood P, McGovern M, Colloca L, King H, van Dorsten B, Pun T, Cheung M. Comparative effectiveness of cognitive behavioral therapy for chronic pain and chronic pain self-management within the context of voluntary patient-centered prescription opioid tapering: the EMPOWER study protocol. Pain Med 2020;21:1523–31.
[22] Davitte JM, Stott-Miller M, Ehm MG, Cunnington MC, Reynolds RF. Integration of real-world data and genetics to support target identification and validation. Clin Pharmacol Ther 2022;111:63–76.
[23] Delorme J, Bertin C, Delage N, Eschalier A, Ardid D, Authier N, Chenaf C. Prevalence of chronic pain in opioid-maintained patients using the capture-recapture method: a nationwide population-based study. PAIN 2021;162:195–202.
[24] Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. PAIN 2005;113:9–19.
[25] Eichler HG, Pignatti F, Schwarzer-Daum B, Hidalgo-Simon A, Eichler I, Arlett P, Humphreys A, Vamvakas S, Brun N, Rasi G. Randomized controlled trials versus real world evidence: neither magic nor myth. Clin Pharmacol Ther 2021;109:1212–8.
[26] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet (London, England) 2007;370:1453–7.
[27] European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making, 2023:289699.
[28] Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. Clin Pharmacol Ther 2019;105:867–77.
[29] Franklin JM, Liaw KL, Iyasu S, Critchlow CW, Dreyer NA. Real-world evidence to support regulatory decision making: new or expanded medical product indications. Pharmacoepidemiol Drug Saf 2021;30:685–93.
[30] Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, Pawar A, Bessette LG, Lee H, Garry EM, Gautam N, Schneeweiss S. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results rrom the RCT DUPLICATE initiative. Circulation 2021;143:1002–13.
[31] Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, Schneeweiss S. Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. Clin Pharmacol Ther 2020;107:817–26.
[32] Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? Clin Pharmacol Ther 2017;102:924–33.

[33] Gerbershagen HJ, Pogatzki-Zahn E, Aduckathil S, Peelen LM, Kappen TH, van Wijck AJM, Kalkman CJ, Meissner W. Procedure-specific risk factor analysis for the development of severe postoperative pain. Anesthesiology 2014;120:1237–45.
[34] Gillman A, Zhang D, Jarquin S, Karp JF, Jeong J-H, Wasan AD. Comparative effectiveness of embedded mental health services in pain management clinics vs standard care. Pain Med 2020;21:978–91.
[35] Gini R, Sturkenboom MCJ, Sultana J, Cave A, Landi A, Pacurariu A, Roberto G, Schink T, Candore G, Slattery J, Trifirò G. Different strategies to execute multi-database studies for medicines surveillance in real-world setting: a reflection on the European model. Clin Pharmacol Ther 2020; 108:228–35.
[36] Gokhale M, Stürmer T, Buse JB. Real-world evidence: the devil is in the detail. Diabetologia 2020;63:1694–705.
[37] Gray CM, Grimson F, Layton D, Pocock S, Kim J. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. Drug Saf 2020;43:623–33.
[38] Hah JM, Aivaliotis VI, Hettie G, Pirrotta LX, Mackey SC, Nguyen LA. Whole body pain distribution and risk factors for widespread pain among patients presenting with abdominal pain: a retrospective cohort study. Pain Ther 2022;11:683–99.
[39] Hak DJ, Mackowiak JI, Irwin DE, Aldridge ML, Mack CD. Real-world evidence: a review of real-world data sources used in orthopaedic research. J Orthop Trauma 2021;35:S6–12.
[40] Harle CA, Listhaus A, Covarrubias CM, Schmidt SO, Mackey S, Carek PJ, Fillingim RB, Hurley RW. Overcoming barriers to implementing patient-reported outcomes in an electronic health record: a case report. J Am Med Inform Assoc 2016;23:74–9.
[41] Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. CHANCE 2019;32:42–9.
[42] Hoang NS, Hwang W, Katz DA, Mackey SC, Hofmann LV. Electronic patient-reported outcomes: semi-automated data collection in the interventional radiology clinic. J Am Coll Radiol 2019;16:472–7.
[43] Hohenschurz-Schmidt D, Kleykamp BA, Draper-Rodi J, Vollert J, Chan J, Ferguson M, McNicol E, Phalip J, Evans SR, Turk DC, Dworkin RH, Rice ASC. Pragmatic trials of pain therapies: a systematic review of methods. PAIN 2022;163:21–46.
[44] Hohenschurz-Schmidt DJ, Cherkin D, Rice ASC, Dworkin RH, Turk DC, McDermott MP, Bair MJ, DeBar LL, Edwards RR, Farrar JT, Kerns RD, Markman JD, Rowbotham MC, Sherman KJ, Wasan AD, Cowan P, Desjardins P, Ferguson M, Freeman R, Gewandter JS, Gilron I, Grol-Prokopczyk H, Hertz SH, Iyengar S, Kamp C, Karp BI, Kleykamp BA, Loeser JD, Mackey S, Malamut R, McNicol E, Patel KV, Sandbrink F, Schmader K, Simon L, Steiner DJ, Veasley C, Vollert J. Research objectives and general considerations for pragmatic clinical trials of pain treatments: IMMPACT statement. PAIN 2023;164:1457–72.
[45] Hwang TJ, Sinha MS, Dave CV, Kesselheim AS. Prescription opioid epidemic and trends in the clinical development of new pain medications. Mayo Clinic Proc 2019;94:2437–43.
[46] Irving E, van den Bor R, Welsing P, Walsh V, Alfonso-Cristancho R, Harvey C, Garman N, Grobbee DE. Series: pragmatic trials and real world evidence: paper 7. Safety, quality and monitoring. J Clin Epidemiol 2017;91:6–12.
[47] Kalkman S, van Thiel GJMW, Zuidgeest MGP, Goetz I, Pfeiffer BM, Grobbee DE, van Delden JJM. Series: pragmatic trials and real world evidence: paper 4. Informed consent. J Clin Epidemiol 2017;89:181–7.
[48] Kerckhove N, Bornier N, Mulliez A, Elyn A, Teixeira S, Authier N, Bertin C, Chenaf C. Prevalence of chronic pain among people with dementia: a nationwide study using French administrative data. Am J Geriatr Psychiatry 2023;31:1149–63.
[49] Kerr NL. HARKing: hypothesizing after the results are known. Pers Soc Psychol Rev 1998;2:196–217.
[50] Komann M, Baumbach P, Stamer UM, Weinmann C, Arnold C, Pogatzki-Zahn E, Meißner W. Desire to receive more pain treatment—a relevant patient-reported outcome measure to assess quality of post-operative pain management? Results from 79,996 patients enrolled in the Pain Registry QUIPS from 2016 to 2019. J Pain 2021;22:730–8.
[51] Li I, Pan J, Goldwasser J, Verma N, Wong WP, Nuzumlalı MY, Rosand B, Li Y, Zhang M, Chang D, Taylor RA, Krumholz HM, Radev D. Neural natural language processing for unstructured data in electronic health records: a review. Computer Sci Rev 2022;46:100511.
[52] Loke YK, Mattishent K. Propensity score methods in real-world epidemiology: a practical guide for first-time users. Diabetes Obes Metab 2020;22:13–20.
[53] Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015; 350:h2147.
[54] Ma C, Wang X, Wu J, Cheng X, Xia L, Xue F, Qiu L. Real-world big-data studies in laboratory medicine: current status, application, and future considerations. Clin Biochem 2020;84:21–30.
[55] Macfarlane GJ, Forget P, Price C, Meissner W, Zaslansky R. Perspectives on pain registries. PAIN 2021;162:2201–3.
[56] Maher DP, Wong CH, Siah KW, Lo AW. Estimates of probabilities of successful development of pain medications: an analysis of pharmaceutical clinical development programs from 2000 to 2020. Anesthesiology 2022;137:243–51.
[57] Makady A, de Boer A, Hillege H, Klungel O, Goettsch W, on behalf of GetReal Work Package 1. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health 2017;20: 858–65.
[58] Meinecke A-K, Welsing P, Kafatos G, Burke D, Trelle S, Kubin M, Nachbaur G, Egger M, Zuidgeest M. Series: pragmatic trials and real world evidence: paper 8. Data collection and management. J Clin Epidemiol 2017;91:13–22.
[59] Monti S, Grosso V, Todoerti M, Caporali R. Randomized controlled trials and real-world data: differences and similarities to untangle literature data. Rheumatology (Oxford) 2018;57(suppl 7):vii54–58.
[60] Nature welcomes registered reports. Nature 2023;614:594.
[61] O'Connell NE, Belton J, Crombez G, Eccleston C, Fisher E, Ferraro MC, Hood A, Keefe F, Knaggs R, Norris E, Palermo TM, Pickering G, Pogatzki-Zahn E, Rice AS, Richards G, Segelcke D, Smart KM, Soliman N, Stewart G, Tölle T, Turk D, Vollert J, Wainwright E, Wilkinson J, Williams AcdC. Enhancing the trustworthiness of pain research: a call to action. J Pain. 2024;104736. doi:10.1038/d41586-023-00506-2.
[62] O'Leary CP, Cavender MA. Emerging opportunities to harness real world data: an introduction to data sources, concepts, and applications. Diab Obes Metab 2020;22(suppl 3):3–12.
[63] Orsini LS, Berger M, Crown W, Daniel G, Eichler H-G, Goettsch W, Graff J, Guerino J, Jonsson P, Lederer NM, Monz B, Mullins CD, Schneeweiss S, Brunt DV, Wang SV, Willke RJ. Improving transparency to build trust in real-world secondary data studies for hypothesis testing-why, what, and how: recommendations and a road map from the real-world evidence transparency initiative. Value Health 2020;23:1128–36.
[64] Oude Rengerink K, Kalkman S, Collier S, Ciaglia A, Worsley SD, Lightbourne A, Eckert L, Groenwold RHH, Grobbee DE, Irving EA. Series: pragmatic trials and real world evidence: paper 3. Patient selection challenges and consequences. J Clin Epidemiol 2017;89:173–80.
[65] Patorno E, Schneeweiss S, Wang SV. Transparency in real-world evidence (RWE) studies to build confidence for decision-making: reporting RWE research in diabetes. Diab Obes Metab 2020;22 (suppl 3):45–59.
[66] Pearl J. Causal inference from indirect experiments. Artif Intelligence Med 1995;7:561–82.
[67] Petrou PA, Leong MS, Mackey SC, Salmasi V. Stanford pragmatiec effectiveness comparison (SPEC) protocol: comparing long-term effectiveness of high-frequency and burst spinal cord stimulation in real-world application. Contemp Clin Trials 2021;103:106324.
[68] Ramamoorthy A, Huang SM. What does it take to transform real-world data into real-world evidence? Clin Pharmacol Ther 2019;106:10–8.
[69] Ramsey SD, Adamson BJ, Wang X, Bargo D, Baxi SS, Ghosh S, Meropol NJ. Using electronic health record data to identify comparator populations for comparative effectiveness research. J Med Econ 2020; 23:1618–22.
[70] Rubinger L, Ekhtiari S, Gazendam A, Bhandari M. Registries: big data, bigger problems? Injury 2023;54(suppl 3):S39–42.
[71] Rudrapatna VA, Butte AJ. Opportunities and challenges in using real-world data for health care. J Clin Invest 2020;130:565–74.
[72] Ryan TA. Multiple comparison in psychological research. Psychol Bull 1959;56:26–47.
[73] Salmasi V, Terkawi AS, Mackey SC. Pragmatic comparative effectiveness trials and learning health systems in pain medicine: opportunities and challenges. Anesthesiol Clin 2023;41:503–17.
[74] Saluan P, Styron J, Ackley JF, Prinzbach A, Billow D. Injury types and incidence rates in precollegiate female gymnasts: a 21-year experience at a single training facility. Orthop J Sports Med 2015;3: 2325967115577596.
[75] Singh G, Schulthess D, Hughes N, Vannieuwenhuyse B, Kalra D. Real world big data for clinical research and drug development. Drug Discov Today 2018;23:652–60.
[76] Soderberg CK, Errington TM, Schiavone SR, Bottesini J, Thorn FS, Vazire S, Esterling KM, Nosek BA. Initial evidence of research quality of registered reports compared with the standard publishing model. Nat Hum Behav 2021;5:990–7.

[77] Spitzer E, Cannon CP, Serruys PW. Should real-world evidence be incorporated into regulatory approvals? Expert Opin Drug Saf 2018;17:1155–9.
[78] Swift B, Jain L, White C, Chandrasekaran V, Bhandari A, Hughes DA, Jadhav PR. Innovation at the intersection of clinical trials and real-world data science to advance patient care. Clin Translat Sci 2018;11:450–60.
[79] Toh S. Analytic and data sharing options in real-world multidatabase studies of comparative effectiveness and safety of medical products. Clin Pharmacol Ther 2020;107:834–42.
[80] Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS, Hewitt DJ, Jadad AR, Katz NP, Kramer LD, Manning DC, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robinson JP, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Witter J. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. PAIN 2003;106:337–45.
[81] Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. Lancet 2011;377:2226–35.
[82] Use of real-world evidence to support regulatory decision-making for medical devices. FDA; 2023. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices. Accessed March 19, 2024.
[83] VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-Value. Ann Int Med 2017;167:268–74.
[84] Vollert J, Kleykamp BA, Farrar JT, Gilron I, Hohenschurz-Schmidt D, Kerns RD, Mackey S, Markman JD, McDermott MP, Rice ASC, Turk DC, Wasan AD, Dworkin RH. Real-world data and evidence in pain research: a qualitative systematic review of methods in current practice. PAIN Rep 2023;8:e1057.
[85] Vollert J, Segelcke D, Weinmann C, Schnabel K, Fuchtmann F, Rosenberger DC, Komann M, Maessen T, Sauer L, Kalso E, Fletcher D, Lavand'homme P, Kaiser U, Liedgens H, Meissner W, Pogatzki-Zahn EM. Responsiveness of multiple patient-reported outcome measures for acute postsurgical pain: primary results from the international multicentre PROMPT NIT-1 study. Br J Anaesth 2024;132:96–106.
[86] Wang SV, Pottegård A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, Berger ML, Crane G, Goettsch W, Hua W, Kabadi S, Kern DM, Kurz X, Langan S, Nonaka T, Orsini L, Perez-Gutthann S, Pinheiro S, Pratt N, Schneeweiss S, Toussi M, Williams RJ. Harmonized protocol template to enhance reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: a good practices report of a joint ISPE/ISPOR task force. Pharmacoepidemiol Drug Saf 2023;32:44–55.
[87] Wang SV, Schneeweiss S, Franklin JM, Desai RJ, Feldman W, Garry EM, Glynn RJ, Lin KJ, Paik J, Patorno E, Suissa S, D'Andrea E, Jawaid D, Lee H, Pawar A, Sreedhara SK, Tesfaye H, Bessette LG, Zabotka L, Lee SB, Gautam N, York C, Zakoul H, Concato J, Martin D, Paraoan D, Quinto K. Emulation of randomized clinical trials with nonrandomized database analyses: results of 32 clinical trials. JAMA 2023;329:1376–85.
[88] Wang SV, Schneeweiss S, Gagne JJ, Evers T, Gerlinger C, Desai R, Najafzadeh M. Using real-world data to extrapolate evidence from randomized controlled trials. Clin Pharmacol Ther 2019;105:1156–63.
[89] Wasan AD, O'Connell B, DeSensi R, Bernstein C, Pickle E, Zemaitis M, Levy O, Jeong J-H, Cooper GF, Douaihy A. The comparative effectiveness of medicinal cannabis for chronic pain versus prescription medication treatment. PAIN 2025;166:1631–40.
[90] Welsing PM, Oude Rengerink K, Collier S, Eckert L, van Smeden M, Ciaglia A, Nachbaur G, Trelle S, Taylor AJ, Egger M, Goetz I. Series: pragmatic trials and real world evidence: paper 6. Outcome measures in the real world. J Clin Epidemiol 2017;90:99–107.
[91] Worsley SD, Oude Rengerink K, Irving E, Lejeune S, Mol K, Collier S, Groenwold RHH, Enters-Weijnen C, Egger M, Rhodes T. Series: pragmatic trials and real world evidence: paper 2. Setting, sites, and investigator selection. J Clin Epidemiol 2017;88:14–20.
[92] Xie J, Wu EQ, Wang S, Cheng T, Zhou Z, Zhong J, Liu L. Real-world data for healthcare research in China: call for actions. Value Health Reg Issues 2022;27:72–81.
[93] Yuan H, Ali MS, Brouwer ES, Girman CJ, Guo JJ, Lund JL, Patorno E, Slaughter JL, Wen X, Bennett D. Real-World evidence: what it is and what it can tell us according to the international society for pharmacoepidemiology (ISPE) comparative effectiveness research (CER) special interest group (SIG). Clin Pharmacol Ther 2018;104:239–41.
[94] Zaslansky R. Status quo of pain-related patient-reported outcomes and perioperative pain management in 10,415 patients from 10 countries: analysis of registry data. Eur J Pain 2022;26:2120–40.
[95] Zaslansky R, Rothaug J, Chapman CR, Bäckström R, Brill S, Fletcher D, Fodor L, Gordon DB, Komann M, Konrad C, Leykin Y, Pogatski-Zahn E, Puig MM, Rawal N, Ullrich K, Volk T, Meissner W. PAIN OUT: the making of an international acute pain registry. Eur J Pain 2015;19:490–502.
[96] Zhang X, Stamey JD, Mathur MB. Assessing the impact of unmeasured confounders for credible and reliable real-world evidence. Pharmacoepidemiol Drug Saf 2020;29:1219–27.
[97] Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE. Series: pragmatic trials and real world evidence: paper 1. Introduction. J Clin Epidemiol 2017;88:7–13.
[98] Zuidgeest MGP, Welsing PMJ, van Thiel GJMW, Ciaglia A, Alfonso-Cristancho R, Eckert L, Eijkemans MJC, Egger M. Series: pragmatic trials and real world evidence: paper 5. Usual care and real life comparators. J Clin Epidemiol 2017;90:92–8.