ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Policy Perspective

Acceptability of Using Real-World Data to Estimate Relative Treatment Effects in Health Technology Assessments: Barriers and Future Steps

Manuel Gomes, PhD, Alex J. Turner, PhD, Cormac Sammon, PhD, Dalia Dawoud, PhD, Sreeram Ramagopalan, PhD, Alex Simpson, MSc, Uwe Siebert, PhD

ABSTRACT

Objectives: Evidence about the comparative effects of new treatments is typically collected in randomized controlled trials (RCTs). In some instances, RCTs are not possible, or their value is limited by an inability to capture treatment effects over the longer term or in all relevant population subgroups. In these cases, nonrandomized studies (NRS) using real-world data (RWD) are increasingly used to complement trial evidence on treatment effects for health technology assessment (HTA). However, there have been concerns over a lack of acceptability of this evidence by HTA agencies. This article aims to identify the barriers to the acceptance of NRS and steps that may facilitate increases in the acceptability of NRS in the future.

Methods: Opinions of the authorship team based on their experience in real-world evidence research in academic, HTA, and industry settings, supported by a critical assessment of existing studies.

Results: Barriers were identified that are applicable to key stakeholder groups, including HTA agencies (eg, the lack of comprehensive methodological guidelines for using RWD), evidence generators (eg, avoidable deviations from best practices), and external stakeholders (eg, data controllers providing timely access to high-quality RWD). Future steps that may facilitate future acceptability of NRS include improvements in the quality, integration, and accessibility of RWD, wider use of demonstration projects to highlight the value and applicability of nonrandomized designs, living, and more detailed HTA guidelines, and improvements in HTA infrastructure relating to RWD.

Conclusion: NRS can represent a crucial source of evidence on treatment effects for use in HTA when RCT evidence is limited.

Keywords: causal inference, comparative effectiveness, health technology assessment, nonrandomized studies, real-world data.

VALUE HEALTH. 2024; 27(5):623–632

Highlights

- Multiple studies have elicited the views of health technology assessment (HTA) agencies and other key stakeholders on barriers to the acceptance of real-world data (RWD) studies more generally, including concerns relating to methodology (eg, residual biases), data availability, governance and quality, and limited trust in RWD studies, as well as barriers relating to HTA policy and partnerships.
- We focus specifically on the barriers to the acceptance of nonrandomized studies (NRS) estimating comparative effectiveness. We outline a pathway determining the acceptance and assign the responsibility for specific barriers to key stakeholders, including (1) evidence-appraisers (HTA agencies), such as providing comprehensive methodological guidelines for conducting NRS, (2) evidence generators (manufacturers), such as whether latest best practices are adhered too, and (3) external stakeholders (eg, data controllers), such as whether there is timely access to high-quality RWD.
- We discuss future steps requiring a joint effort between all stakeholders that could facilitate increased acceptability of NRS, including improvements in the quality and accessibility of RWD, use of demonstration projects to benchmark NRS results against RCT evidence and highlight the value and applicability best-practice methods, and ensuring up-to-date and detailed HTA real-world evidence guidelines.

Introduction

Randomized controlled trials (RCTs) are widely considered to be the gold-standard for assessing the comparative effectiveness and safety of health technologies. However, there are situations where RCTs are not possible, for example, in rare diseases in which it is difficult to design well-powered RCTs. In these cases, nonrandomized studies (NRS) using (partly) routinely collected "real-world data" (RWD) in administrative data sets, disease registries, and electronic health records may offer the best alternative approach to estimating comparative effectiveness. These NRS can also complement randomized evidence, by providing estimates of treatment effects in a routine care setting, over the longer term, or in population subgroups typically excluded from RCTs.

Regulatory drug approvals have been increasingly informed by NRS, particularly for treatments of serious and life-threating conditions.[1,2] Consideration of evidence from real-world external control arms (RW-ECAs) providing comparator data for single-arm trials (SATs) in submissions to the US Food and Drug Administration (FDA) have been cited as a key driver of the rapid development of cancer drugs.[3] The use of NRS for regulatory approval has subsequently affected the evidence available at the time of health technology assessment (HTA), with NRS (particularly from RW-ECAs) becoming increasingly prevalent in HTA submissions.[4,5]

Other uses of RWD, for example, for identifying unmet need, assessing the generalizability of RCT populations, extrapolating trial outcomes, and for deriving inputs to economic models other than comparative effectiveness (eg, resource use, utilities, event rates, and disease prevalence and incidence), have been broadly accepted for informing reimbursement decisions.[6] However, concerns around the limitations of NRS, most importantly the threat of residual confounding due to the absence of randomization and other pitfalls, such as information and selection biases, have discouraged the full acceptance of these studies as substitutes for RCTs by HTA agencies,[7] with a recent survey of EUnetHTA member HTA organizations highlighting that acceptance is lowest for combination therapies and personalized medicine.[8] For example, the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany have explicitly raised concerns over NRS replacing RCTs[9] and often have not considered NRS evidence even if of high quality and used by other HTA agencies to inform decisions. HTA agencies' reluctance toward NRS has persisted despite recent advancements in best practices relating to study design, data source selection, the application of advanced causal inference methods, and the increased use of quantitative bias analysis (QBA) to systematically explore the impact of residual biases.

Multiple studies have elicited the views of HTA agencies and other key stakeholders on barriers to the acceptance of RWD studies more generally,[10–14] including concerns relating to methodology (eg, residual biases), data availability, governance and quality, and limited trust in RWD studies, as well as barriers relating to HTA policy and partnerships.[15] In this policy perspective, we focus specifically on the barriers to the acceptance of NRS in HTA, that is, barriers preventing the routine use of NRS for decision making when RCT evidence is absent or limited. We also outline a pathway determining the acceptance of NRS and assign the responsibility for specific barriers to key stakeholders, including (1) evidence-appraisers (HTA agencies), such as providing comprehensive methodological guidelines for conducting NRS, (2) evidence generators (manufacturers), such as whether latest best practices are adhered too, and (3) external stakeholders (eg, data controllers), such as whether there is timely access to high-quality RWD.

We also discuss future steps requiring a joint effort between all stakeholders that could facilitate increased acceptability of NRS, including improvements in the quality and accessibility of RWD, use of demonstration projects to benchmark NRS results against RCT evidence and highlight the value and applicability best-practice methods, and ensuring up-to-date and detailed HTA real-world evidence (RWE) guidelines. Such steps have the potential to harmonize the generation and acceptance of high-quality NRS, which could enable faster and better-informed healthcare resource allocation and decision making where RCT evidence is not available.

These barriers and solutions primarily relate to the acceptance of NRS at initial HTA submission (in which RW-ECAs will be commonplace), but equally apply to NRS conducted for comparative effectiveness reappraisal (where RWD might be the sole data source), which is increasingly cited as an important component of life-cycle approaches to HTA.[16–18] These barriers and solutions reflect the opinions of the authorship team based on their experience in RWE research in academic, HTA agency, and industry settings and supported by evidence from a critical assessment of the existing literature. The views are those of the authors alone and do not necessarily reflect the views of their employers.

Barriers to the Acceptance of NRS

Figure 1 outlines a conceptual framework determining the acceptance of NRS by HTA agencies. It demonstrates that NRS will only be accepted by HTA agencies if 4 requirements are met. First, that requirements and procedures in methodological RWE guidelines published by HTA agencies are consistent with best practices. Second, that it is possible to implement best practices outlined in these guidelines, which is partly determined by whether data controllers produce high-quality RWD that can be accessed in a timely manner. Third, that best practices are implemented by manufacturers (or by external practitioners conducting analysis on their behalf) appropriately and that implementation of best practices can produce evidence of sufficient quality to support HTA decision making. This is in part determined by whether causal inference methods (developed by expert methodologists) are available to address these needs. Finally, acceptance requires that, conditional on high-quality NRS being included in HTA submissions, HTA agencies have the tools and knowledge in place to adequately judge quality and accept this evidence. In the following, we discuss each of these requirements.

Limitations of guidelines

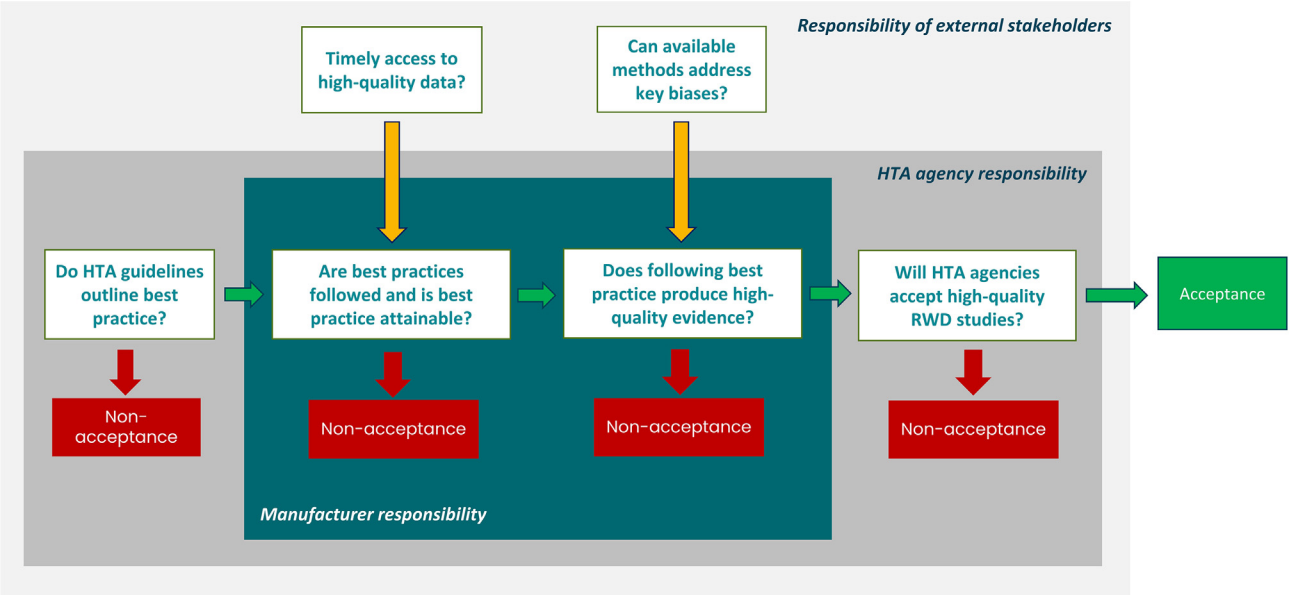
Best practices relating to NRS describe methods and approaches that minimize key biases in the use of RWD. Biases primarily include threats to internal validity, such as confounding, information bias, selection bias, time-related biases (eg, immortal time bias), and reverse causation.[19,20] However, when NRS use data from outside of a decision maker's target population (eg, data from another jurisdiction or care setting), biases relating to threats to external validity are also relevant and can be caused by characteristics driving treatment effect heterogeneity (effect modifiers) differing between study populations (estimation samples) and the target population.[21] Concerns over the lack of "transportability" of treatment effects have implicitly led to a preference for the use of local RWD.[22]

Figure 2 provides a summary of these best practices, based on the latest published literature. This includes the systematic identification and evaluation of alternative RWD sources, the use of target trial emulation (TTE) in study design,[23,24] use of causal inference methods, such as matching, weighting, and g-methods for minimizing observed baseline and time-varying confounding,[25] quasi-experimental methods for addressing unobserved confounding,[25] and methods for minimizing information bias[26] and bias due to a lack of external validity.[27,28] The figure also describes recommendations for sensitivity analysis to explore the robustness of results to residual bias,[29–32] for the clear reporting of results,[20,33,34] and for the selection of appropriate RWD sources.[35]

However, unlike methodological guidance relating to RCTs, best practices are rarely comprehensively described in methodological guidelines published by HTA agencies. A review of 41 published guidelines from regulators, HTA agencies, RWD initiatives, and professional societies found that HTA RWE guidelines failed to provide sufficient guidance on criteria relating to study design, data quality and relevance, analytical methods, and transparency/reproducibility.[36] This incompleteness is found in international methodological guidelines too, meaning that HTA agencies relying on references to other guidelines for more detail (as done in guidelines from Haute Autorité de Santé) will also be providing incomplete information.

In cases where key elements of best practice are absent from guidelines, those carrying out NRS must determine best practices using fragmented information from other sources. In some cases, approaches taken will diverge from those expected by HTA committees, especially in cases where there is limited consensus across guidelines about what constitutes best practice. For example, a recent press release from IQWiG criticized submissions

Figure 1. Barriers to the acceptance of NRS by HTA agencies.



HTA indicates health technology assessment; NRS, nonrandomized study; RWD, real-world data.

for not conducting sensitivity analysis,[37] but IQWiG guidelines provide little detail on what specific sensitivity analyses are recommended and fail to describe methods for key biases other than confounding.[22]

However, the recently published National Institute for Health and Care Excellence (NICE) RWE framework represents a crucial step forward.[38] The framework provides detailed recommendations for comparative effectiveness studies relating to assessing data suitability, study design (TTE), analysis to minimize multiple sources of bias, sensitivity analysis to explore residual biases (including how QBA can be used alongside other approaches), and reporting standards. However, gaps still exist. For example, although the framework states that a "transparent, systematic, and reproducible process" should be used to identify confounders using published literature and expert opinion, it does not highlight tools that can be used elicit expert opinion, such as advisory boards or consensus methods, such as Delphi methods. Because these tools are not specifically used only for NRS, such methods need not be described in detail, but reference to external information on these tools are also absent. Similarly, as with other HTA RWE guidelines, the framework does not highlight methods that can increase transportability and fails to outline all methods that can be used to adjust for confounding (including doubly robust methods). Although the NICE Decision Support Unit has published technical support documents relating to NRS methods,[39,40] the latest was released in 2016 and does not cover the current best practices relating to primary and sensitivity analysis.

Challenges with following best practice, data quality, and access

In cases where HTA RWE guidelines provide detailed information on study approaches, it is then the responsibility of manufacturers to ensure these best practices are followed by studies informing HTA. If best practices are not followed, the quality of evidence may be questionable and may be rejected by HTA agencies. Although comprehensive HTA RWE guidelines will increase adherence to best practice, the absence of these should not be used to justify deviations from best practice. It is the ultimate responsibility of manufacturers to ensure analysis is robust, and analysts should follow methodological guidelines when conducting NRS.

Findings from multiple studies suggest that avoidable deviations from best practices are common. A review of NRS identified that 81% of studies had at least 1 major methodological issue known to cause substantial bias.[19] Similarly, a recent press release from IQWiG outlined frequent nonadherence to their guidelines.[37] Consequences of such deviations can be severe, with biases of ~50% in either direction if rules of TTE and adjustment for time-varying confounding are ignored.[41] Also, in response to submissions using RW-ECAs as pivotal evidence, NICE have criticized submissions for not systematically identifying and evaluating RWD sources, leaving uncertainty over whether alternative RWD sources could have been more suitable.[42,43]

However, manufacturers following best practices in design and analysis may not be sufficient if RWD is not of high quality and/or is not readily available, and this is arguably the responsibility of external stakeholders rather than manufacturers. For example, a failure to control for all relevant confounders or emulate eligibility criteria of the target trial may be driven by insufficient richness of available RWD, an issue cited by HTA agencies as a driver of rejection of NRS[37] and as barrier to the adoption of RWD more generally.[8]

In addition, RWD may often be used to estimate comparative effectiveness when timely decisions are needed on life-saving therapies.[44-46] Gaining access to detailed RWD, including those collected and funded by public bodies, is frequently a costly and lengthy process. In some instances, RWD studies may face a trade-off between the richness of the data and a need to ensure timeliness of decision making.

Furthermore, many HTA agencies state a preference for the use of local data[22,38] and report low acceptance of RWD from other countries.[8,37] However, the availability of high-quality RWD varies considerably across countries, meaning that the choice of RWD source often involves a compromise between data quality and locality.

Figure 2. Summary of best practices in the design and conduct of NRS.

NRS indicates nonrandomized study.

Finally, when examining effectiveness in rare diseases/populations, although larger samples are a potential strength of using RWD over RCTs, sample sizes can still be small (particularly when RW-ECA designs are used). This may drive manufacturers to control for only a subset of confounders, despite data on other confounders being available.

Of course, it is sensible for HTA agencies to consider bias resulting from deviations from best practice as a limitation, irrespective of the reason for these deviations. However, where deviations are out of the control of manufacturers, there may be cases where HTA agencies should recognize that the "best available evidence" in certain therapeutic/disease areas will be associated with greater levels of uncertainty and consider adjusting the expected level of certainty needed to inform decision making in these cases.

The potential for unexplained biases

Conditional on best practices being followed and high-quality RWD being available, acceptance of NRS to inform HTA decision making depends on the extent to which unexplained biases are possible.

Unlike RCT evidence, consistency of estimates from NRS will always rely on untestable assumptions that are open to criticism. For example, even if data are available on all confounders identified via experts and/or literature reviews, gaps in expert knowledge and the existence of unmeasurable confounders may lead to residual confounding.

Although quasi-experimental designs can help address unobserved confounding, they also require distinct untestable assumptions to be met. For example, the validity of commonly used geographical variables used in instrumental variable designs have been shown to be questionable[47] and have led IQWiG to highlight potential limitations of instrumental variable methods for comparative effectiveness purposes.[22]

Many studies have attempted to assess residual bias in NRS by examining their ability to reproduce results from RCTs. These have shown that self-inflicted biases are often a bigger problem than unmeasured confounding by demonstrating that when using appropriate study design, causal frameworks and appropriate analytical methods, NRS can reliably approximate results from well-conducted RCTs.[41,48-53] However, such replication is not always successful.[54] Although early results from RCT DUPLICATE suggest success in replication, full emulation was difficult for some RCTs, and the study identified that differences between RCT-derived and NRS-derived effects would have affected regulatory conclusions in 40% of studies.[55] Difficulty in replication tends to be largest for placebo-controlled RCTs because of a higher risk of intractable confounding because in clinical practice nonusers are likely those with very mild disease (requiring no treatment) or very severe disease (with contraindication for all treatments).[55,56]

The potential for greater residual bias in some settings means that acceptance of NRS by HTA agencies will likely continue to be made on a case-by-case basis. We note however that QBA methods allow the impact of residual biases to be systematically explored, by examining the size of bias required to change a result (eg, treatment effect relative to some decision-making threshold) or estimating the direction, magnitude, and uncertainty of bias associated with treatment effects. Although QBA methods can only reduce uncertainty with evidence and not eliminate it, we argue that the use of QBA should mean that high-quality NRS should be accepted by HTA because it allows uncertainty to be clearly assessed and documented. In these cases, HTA agencies could also make use of conditional reimbursement schemes, where available, making a decision based on NRS evidence while requiring randomized evidence or further NRS evidence be collected.

Ingrained reluctance or inability to accept even high-quality NRS by HTA agencies

A final potential barrier to acceptance is the rejection of high-quality NRS by HTA agencies. HTA agencies are often clear that NRS will only be considered in cases which RCT evidence cannot be produced, and some argue that, in the long-run, RCTs are always possible.[57] If manufacturers are unable to convince HTA committees that an RCT was not possible, then even well-conducted NRS could be rejected. However, we note it is often unclear what metrics HTA committees use to determine the feasibility of an RCT.

Second, acceptance is harmed by ingrained perceptions that NRS are always of low quality. Consistent with quality-assessment tools frequently used in clinical guidelines,[58] consensus among HTA bodies is that the quality of evidence on treatment effects from NRS will always be inferior. For example, IQWiG states that the "accuracy of assumptions and confounding adjustment can never be validated. Therefore, such analyses can only be considered to provide a low degree of qualitative certainty."[22] Although it is sensible for HTA agencies to prefer RCT evidence in cases which they are available, there is a concern that evidence from NRS is being judged identically without sufficient reflections on their quality.

Some HTA agencies or committees may also have insufficient expertise to correctly assess NRS. In cases which guidelines are not precise, acceptability of analysis is more dependent on the knowledge of HTA committees on best practices and understanding of causal methods. Limited expertise relating to NRS may lead to nonrobust NRS being accepted for decision making or risk-averse committees rejecting high-quality evidence. This issue was highlighted in a qualitative study informing the CanREValue Collaboration, in which stakeholders cited that the lack of knowledge of best-practice methods made NRS easy to criticize.[59] In addition, user profiles accompanying the NICE RWE framework state that an aim of the framework is to improve knowledge of HTA committees in NRS methods and indicate that relevant expertise in this area is currently limited.[38] To reduce this concern, manufacturers should ensure best practices are implemented and that any NRS methodological guidelines followed are clearly cited to highlight their use to committees.

Finally, there exist institutional barriers that limit the ability of HTA agencies to accurately assess NRS. The CanREValue Collaboration highlights that some HTA agencies are already stretched beyond capacity in terms of finances, expertise, and leadership. The complexity of NRS means they take longer to evaluate; therefore, adding evaluation of NRS on top of existing workloads may be resulting in limited capacity to conduct detailed assessments.[59] The collaboration also cited a historical lack of engagement with industry partners as a barrier, with key stakeholders working in silos.[59]

Future Steps to Increase the Acceptability of Nonrandomized Evidence

Potential steps to reduce barriers to the acceptance of NRS by HTA agencies are outlined in Figure 3.

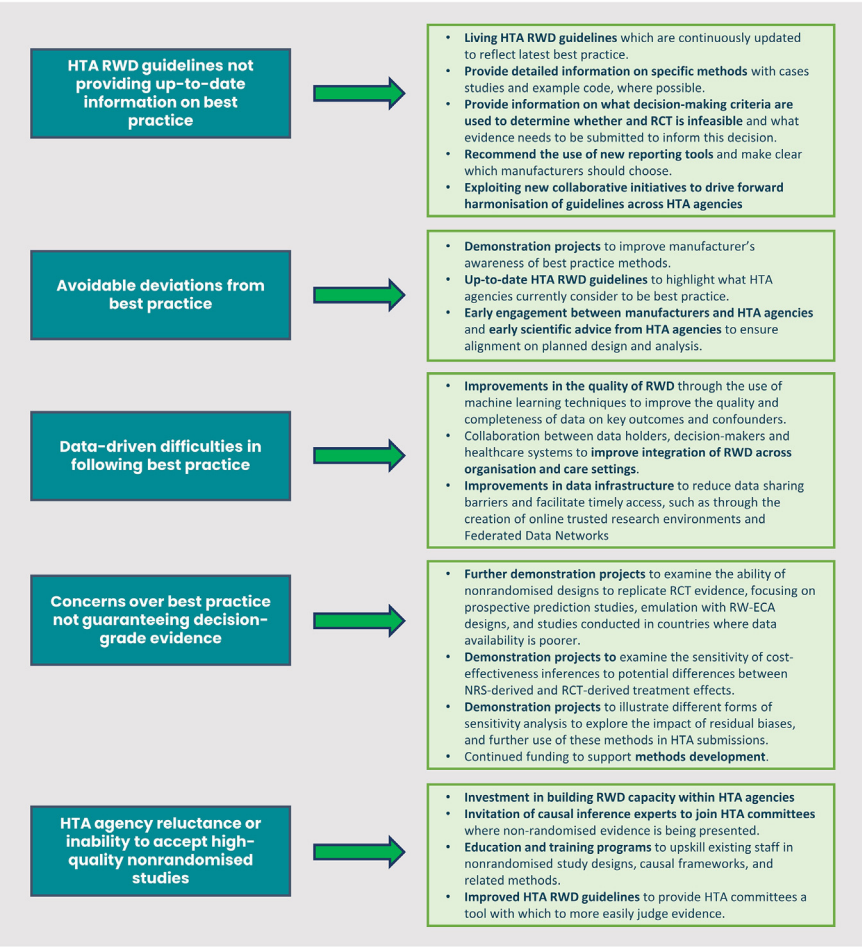
Improvements in HTA RWE guidelines

If NRS are to be conducted as rigorously as RCTs, gaps in HTA RWE guidelines need to be addressed and the level of guidance needs to mirror that for RCTs. For example, the NICE Decision Support Unit publishes technical support documents (TSDs) provide detailed guidance on the use of analytical methods for HTA studies, along with case studies and example code.[60] However, such TSDs have focused on guidance for both RCT-based analyses or decision modeling informing HTA. A TSD to accompany the NICE RWE framework is warranted, or existing TSDs covering NRS[39,40] should be updated to include latest best practices, including guidance on rigorous study design (eg, TTE), contemporary causal inference methods (eg, g-methods for time-varying confounding), and sensitivity analysis to address potential residual confounding (eg, QBA). Other HTA agencies should consider providing guidance of similar detail. More detailed guidance not only provides clarity on what HTA agencies expect and will accept, they can also provide important considerations regarding the application of methods specifically in a HTA context.

Furthermore, HTA guidelines should be clear on what decision-making criteria they are using to determine whether an RCT is not feasible and suggest what evidence manufacturers need to produce to inform this decision.

In addition, although guidelines may need to differ across countries because of differences in decision criteria, greater harmonization across decision makers is warranted. Although numerous payer-led or payer-involved collaborative initiatives (such as RWE4Decisions) have been summarized previously,[15] new collaborative initiatives have been announced and should be exploited to drive further harmonization both across payers and between payers and regulators. The AUS-CAN-NZ-UK

Figure 3. Future steps to reduce the barriers to acceptance of NRS.



HTA indicates health technology assessment; NRS, nonrandomized study; RCT, randomized controlled trial; RWD, real-world data; RW-ECA, real-world external control arm.

collaboration added Institut National d'excellence en Santé et Services Sociaux (INESS) and Pharmac to its collaboration agreement in July 2023, and a joint approach to improving HTA and regulatory collaboration is cited as a key priority.[61] EU HTA regulation was passed in December 2021, introducing a requirement for joint clinical assessment (JCA) of technologies and an option for accessing early advice via joint scientific consultation.[62] Plans for new guidance, created by a new Coordination Group (comprising representatives from national HTA agencies), offers the opportunity to further outline a shared position on NRS across EU HTA agencies. This will build on the latest EUnetHTA guidance, which hinted at more cautious approach to NRS by stating that, because of greater uncertainty surrounding results from NRS, only large NRS-generated treatment effects will be deemed acceptable.[63] JCAs will initially focus on therapies in which NRS features more frequently in submissions (cancer drugs in January 2025, orphan drugs in January 2028); therefore, decisions from these will strengthen our understanding on the EU's latest views on the acceptance of NRS. JCA recommendations are nonbinding; therefore, deviations from joint recommendations by national HTA agencies should be identified because these will act as a barometer for the degree of harmonization. In addition, because joint scientific consultation members also include representatives from the European Medicines Agency (EMA), this process may encourage greater harmonization between the EMA and EU HTA agencies. Joint guidance between regulators and HTA agencies, in situations which are appropriate, should be an end goal.

HTA RWE guidelines must also be living guidelines. Best practices are fast-moving and HTA agencies must keep abreast of the latest methods development to ensure they do not become out-of-date. Other HTA agencies should follow NICE's example, who have committed to continually update their RWE framework.

As highlighted in the NICE RWE framework's user profiles, detailed guidelines may also stimulate removal of other barriers.[38] For example, these can improve knowledge of HTA committees by providing accessible summaries of best practice, act as a tool to connect HTA bodies and manufacturers at an early stage, mitigate potential design and analytical flaws, provide directions for improvements in data quality, and help identify gaps in best practice for which more methodological research might be needed.

In addition to methodological guidance, wider use of reporting tools is crucial for ensuring transparency on behalf of manufactures (and consequently, improving trust in NRS results) and ensuring that HTA agencies have the information they need to correctly judge quality. Tools such as the EUnetHTA's Registry Evaluation and Quality Standards Tool (REQueST) tool[64] and NICE's Data Suitability Assessment Tool (DataSAT) tool[35] should be used to ensure key data source characteristics and justification of

the suitability for answering a specific research question is reported, and the new HARmonized Protocol Template to Enhance Reproducibility (HARPER)[33] and under-development TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guidelines[34] should be considered for reporting key elements of study design and analysis. HTA RWE guidelines should be clear regarding which reporting tools they expect manufacturers to use.

Improvements in the quality, integration, and access to existing RWD sources

An important step to encourage further acceptance of NRS in HTA is to improve the timely access and quality of RWD. Regarding data quality, machine learning (ML) techniques, such as natural language processing, have the potential to increase the efficiency of identifying information on key variables from unstructured data compared with costly and time-consuming manual abstraction without compromising data accuracy.[65] For example, Flatiron Health used natural language processing to derive an algorithm to identify patients with metastatic breast cancer, which demonstrated considerable sensitivity with no differences in average outcomes and baseline characteristics relative to test cohorts generated using only manual abstraction.[66] The breadth of variables included in RWD sources can be increased through integration of multiple existing data sources across organizations and different care settings. Current examples in the United Kingdom include the DATA-CAN initiative, which is a collaboration between industry, academia, and NHS organizations to link oncology data from multiple sources,[67] and the NHS Digital Secure Data Environment, which provides linked de-identified patient-level data across multiple care settings.[68,69]

Improvements in data infrastructure, for example, the creation of digital platforms for accessing data, such as OpenSAFELY and the Secure Data Environment, and the creation of Federated Data Networks, such as the European Health Data and Evidence Network and DARWIN EU,[70] have the potential to facilitate more timely access to RWD. Such infrastructure can also help standardize data management pathways across data sources and overcome data sharing barriers.[71] The use of common data models, such as the Observational Medical Outcomes Partnership (OMOP) CDM, which has been adopted by the EMA,[72] allow standardized analytical packages to be used across multiple international data sources and could stimulate greater acceptance of international RWD. The EU Data Quality Framework and the linked RWD use cases[73] may provide a tool to ensure the quality of data sources and subsequently increase standardization.

Demonstration projects

Demonstration projects are required that focus specifically on methods addressing the most pressing concerns of HTA agencies and describing situations in which they are most appropriate. These could build on existing projects, such as those exploring how reassessments using RWD can inform living guidelines,[16] and those highlighting the usefulness of QBA in a HTA setting.[29,74,75]

First, it is important that the number and type of trial emulation studies continues to expand. Emulation studies have primarily focused on NRS using a single RWD source, and further studies are required to assess whether RCT findings can be replicated using RW-ECA designs that are prevalent in initial HTA submissions (for example, by analyzing RCTs as SATs and replacing randomized control arms with RW-ECAs). Emulation studies have also primarily been conducted using data from the United States and Europe, and studies are needed to examine whether RCTs conducted outside of these regions (where quality and availability of RWD is more limited) can be replicated using NRS. Also, a movement away from studies retrospectively replicating published RCT results (which are subject to researcher bias) to studies prospectively predicting results from ongoing RCTs should continue.[76-78] Finally, given the proliferation of trial emulation studies, systematic literature reviews may be useful for identifying when and where emulation is most successful and therefore when and where the use of NRS is most appropriate.

Second, projects should extend RCT DUPLICATE, which currently examines consequences for regulatory decisions, by examining the degree to which differences between NRS-derived and RCT-derived treatment effects would affect HTA decisions (where cost-effectiveness models may use NRS as inputs). This may involve calculating impacts on incremental cost-effectiveness ratios and on subsequent conclusions regarding cost-effectiveness (ie, whether differences cause incremental cost-effectiveness ratios to shift across common cost-effectiveness thresholds).

Third, as highlighted previously, QBA methods are crucial for encouraging the acceptance of high-quality NRS by allowing the impact of potential residual biases to be systematically explored in a way that facilitates decision making under uncertainty. Further empirical applications of these methods are warranted, particularly in treatments for rare diseases and rare-driver mutations in oncology, for which small RWD samples are likely to result in incomplete adjustment for observed confounding. Similarly, applications of negative control outcomes, cited as an additional approach for examining residual bias in the NICE RWE framework, would be very valuable, especially because identifying these outcomes in RWD might be challenging. In addition, the Observational Patient Evidence for Regulatory Approval Science and Understanding Disease (OPERAND) project showed that 2 different research teams independently emulating RCTs using the same data sources, made different design (eg, eligibility criteria) and analysis (eg, estimation methods) decisions, resulting in different RWD cohorts.[79] This suggests that researcher decision making may be an important source of variation in RWD studies, and further research assessing its impact on estimated treatment effects, is warranted.

Fourth, existing demonstration projects focus primarily on the evaluation of time-fixed treatments, but time-varying/dynamic treatment strategies are increasingly common (eg, in personalized medicine). Building on recent examples,[41,80] applications adjusting for time-varying confounding, and the use of QBA in this setting, are needed.

Fifth, given that concerns over the use of international RWD is cited as an important barrier by HTA agencies, demonstrating the use of methods to increase external validity are crucial, especially given the limited knowledge of these methods by practitioners and the dearth of applications specifically adjusting for both confounding and differences in effect modifiers between study and target populations.[81] Studies applying these methods with RW-ECA designs are particularly needed because the possibility of data from SATs and ECA being drawn from different countries adds additional complexity.

Sixth, recent studies have highlighted the role ML could play in RWD studies,[65] including techniques to identify confounders and functional forms, such as LASSO,[82] and casual inference methods that incorporate the use of ML, including double/debiased ML.[83] However, the limitations of these methods are unknown, and their complexity means that they are less accessible to both practitioners and HTA agencies. Case studies are needed to explore the circumstances in which ML methods can complement existing approaches in meaningful ways.

Lastly, further applications of quasi-experimental methods for estimating cost-effectiveness would be useful, particularly in settings that the assumptions required by these methods are likely to be plausible.[84]

Improvements in HTA infrastructure

Further investments within HTA agencies are needed to improve their capacity to assess the quality of NRS. This could be facilitated by inviting causal inference experts to join HTA committees in which nonrandomized evidence is being presented and by supporting education programs, such as the European Health Data and Evidence Network (EHDEN) academy,[85] to upskill staff in best practices in study design and analytical methods.

Investment is also needed to maintain the rapid development of NRS methods that will increase the possibility that following best practices leads to decision-grade evidence. Continued funding support for RWE initiatives are central to this, with examples including the EU funding of the GetReal Initiative[86] and specific methods development work via the HORIZON Europe programme,[87] and continued regulatory and HTA support (eg, from the FDA and NICE) for RWE/RWD demonstration projects.[88,89] Where HTA agencies and regulators are interested in similar/identical topics, the creation of joint funding sources and forums to share experiences may represent useful ways to stimulate and improve learnings from research activities. Coworking opportunities identified in the section "Improvements in HTA RWE guidelines" may present useful avenues for implementation.

Finally, early engagement between manufacturers and HTA agencies is crucial for ensuring alignment on whether a planned NRS is likely to be appropriate and whether key design and analysis elements need to be adapted to increase acceptability. HTA agencies need to ensure that mechanisms are in place to facilitate the sharing of early scientific advice relating to NRS. Semi-binding commitments on acceptance following from this advice may incentivize detailed discussion and HTA agency investment in skilled staff.[36]

Conclusion

NRS have key limitations for the estimation of treatment effects but may provide crucial evidence for assessing the comparative effectiveness of health technologies in cases that RCT evidence is limited. This study identified key barriers to acceptability of nonrandomized evidence by HTA agencies that could be addressed by the actions of manufacturers, HTA agencies, and external stakeholders. We propose steps for improving the acceptability of NRS in HTA requiring joint effort from all stakeholders, including improvements in guidelines, greater quality and accessibility of RWD, expansion of demonstration projects, and stronger HTA infrastructure.

Author Disclosures

Author disclosure forms can be accessed in the Supplemental Material section.

The views expressed are those of the authors and not those any of the funders.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2024.01.020.

Article and Author Information

Accepted for Publication: January 24, 2024

Published Online: March 25, 2024

doi: https://doi.org/10.1016/j.jval.2024.01.020

Author Affiliations: Department of Applied Health Research, University College London, London, England, UK (Gomes); Putnam PHMR, London, England, UK (Turner, Sammon); Science, Policy and Research Programme, National Institute for Health and Care Excellence, London, England, UK (Dawoud); Faculty of Pharmacy, Cairo University, Cairo, Egypt (Dawoud); Lane Clark and Peacock, London, England, UK (Ramagopalan); Global Access, F. Hoffmann-La Roche Ltd, Grenzacherstrasse, Basel, Switzerland (Simpson); Department of Public Health, Health Services Research and Health Technology Assessment, UMIT TIROL – University for Health Sciences and Technology, Hall in Tirol, Austria (Siebert); Center for Health Decision Science and Department of Health Policy and Management, Harvard T.H Chan School of Public Health, Boston, MA, USA (Siebert); Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA (Siebert).

Correspondence: Alex Simpson, MSc, Global Access, F. Hoffmann-La Roche Ltd, Grenzacherstrasse, Basel, Switzerland. Email: alex.simpson.as1@roche.com

Author Contributions: *Concept and design:* Gomes, Turner, Sammon, Dawoud, Ramagopalan, Simpson, Siebert
Analysis and interpretation of data: Turner, Sammon, Dawoud, Ramagopalan, Simpson
Drafting of the manuscript: Gomes, Turner, Sammon, Siebert
Critical revision of the paper for important intellectual content: Gomes, Dawoud, Ramagopalan, Simpson, Siebert
Drs Gomes and Dawoud are editors for *Value in Health* and had no role in the peer-review process of this article.

Funding/Support: Dr Gomes is recipient of a NIHR Advanced Fellowship (NIHR302259). Authors employed by Putnam PHMR (Drs Turner and Sammon) received financial support from F. Hoffmann-La Roche Ltd to prepare this manuscript. Dr Simpson is an employee of F. Hoffmann-La Roche Ltd and Dr Ramagopalan was an employee of F. Hoffmann-La Roche Ltd during the conduct of this study. Dr Dawoud received support from the European Health Data and Evidence Network project. European Health Data and Evidence Network received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 806968. The Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations. Dr Siebert receives institutional support from the European Union's Horizon 2020 and Horizon Europe programmes in the area of health technology assessment and causal inference.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

REFERENCES

1. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open*. 2016;6(6):e011666.
2. Goring S, Taylor A, Müller K, et al. Characteristics of non-randomised studies using comparisons with external controls submitted for regulatory approval in the USA and Europe: a systematic review. *BMJ Open*. 2019;9(2):e024895.
3. Agrawal S, Arora S, Vallejo JJ, et al. Use of single-arm trials to support malignant hematology and oncology drug and biologic approvals: a 20-year FDA experience. *J Clin Oncol*. 2021;39(suppl 15):e13572.
4. Patel D, Grimson F, Mihaylova E, et al. Use of external comparators for health technology assessment submissions based on single-arm trials. *Value Health*. 2021;24(8):1118–1125.
5. Bullement A, Podkonjak T, Robinson MJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. *Int J Technol Assess*. July 10, 2020. https://doi.org/10.1017/S0266462320000434. published online.
6. Coppo C, Rey L, Carelli A, et al. HTA acceptance of RWE in healthcare decision making: a critical review from seven countries. Roche. https://www.ispor.org/docs/default-source/euro2022/isporeu22thokagevistk.pdf?sfvrsn=18cacb28_0. Accessed August 23, 2023.
7. Value of real-world evidence in health technology assessment lost in translation? Economist Impact. https://cdn.vev.design/private/BCwBc9ZFZyVz8yQQKr9VeLxSnjf1/gZWp0JKu0a_RocheRealWorld_Report_A4_final.pdf. Accessed February 15, 2023.

8. Hogervorst MA, Pontén J, Vreman RA, Mantel-Teeuwisse AK, Goettsch WG. Real world data in health technology assessment of complex health technologies. *Front Pharmacol*. 2022;13:837302.
9. Wieseler B, Neyt M, Kaiser T, Hulstaert F, Windeler J. Replacing RCTs with real world data for regulatory decision making: a self-fulfilling prophecy? *BMJ*. 2023;380:e073100.
10. Sievers H, Joos A, Hiligsmann M. Real-world evidence: perspectives on challenges, value, and alignment of regulatory and national health technology assessment data collection requirements. *Int J Technol Assess Health Care*. 2021;37:e40.
11. Makady A, van Veelen A, Jonsson P, et al. Using real-world data in health technology assessment (HTA) practice: a comparative study of five HTA agencies. *Pharmacoeconomics*. 2018;36(3):359–368.
12. Makady A, Ham RT, de Boer A, et al. Policies for use of real-world data in health technology assessment (HTA): A comparative study of six HTA agencies. *Value Health*. 2017;20(4):520–532.
13. Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU-actions for stakeholders. *Int J Technol Assess Health Care*. September 3, 2020. https://doi.org/10.1017/S026646232000063X. published online.
14. Oortwijn W, Sampietro-Colom L, Trowman R. How to deal with the inevitable: generating real-world data and using real-world evidence for HTA purposes - from theory to action. *Int J Technol Assess Health Care*. 2019;35(4):346–350.
15. Capkun G, Corry S, Dowling O, et al. Can we use existing guidance to support the development of robust real-world evidence for health technology assessment/payer decision-making? *Int J Technol Assess Health Care*. 2022;38(1):e79.
16. Jaksa A, Bloudek L, Carlson JJ, et al. Key learnings from Institute for Clinical and Economic Review's real-world evidence reassessment pilot. *Int J Technol Assess Health Care*. 2022;38(1):e32.
17. Kirwin E, Round J, Bond K, McCabe C. A conceptual framework for life-cycle health technology assessment. *Value Health*. 2022;25(7):1116–1123.
18. Regier DA, Pollard S, McPhail M, et al. A perspective on life-cycle health technology assessment and real-world evidence for precision oncology in Canada. *NPJ Precis Oncol*. 2022;6(1):76.
19. Bykov K, Patorno E, D'Andrea E, et al. Prevalence of avoidable and bias-inflicting methodological pitfalls in real-world studies of medication safety and effectiveness. *Clin Pharmacol Ther*. 2022;111(1):209–217.
20. D'Andrea E, Vinals L, Patorno E, et al. How well can we assess the validity of non-randomised studies of medications? A systematic review of assessment tools. *BMJ Open*. 2021;11(3):e043961.
21. Westreich D, Edwards JK, Lesko CR, Cole SR, Stuart EA. Target validity and the hierarchy of study designs. *Am J Epidemiol*. 2019;188(2):438–443.
22. *Institute for quality and efficiency in health care. Concepts for the generation of routine practice data and their analysis for the benefit assessment of drugs according to §35a Social Code Book V (SGB V)*. Inst Qual Effic Health Care; 2020.
23. Gomes M, Latimer N, Soares M, et al. Target trial emulation for transparent and robust estimation of treatment effects for health technology assessment using real-world data: opportunities and challenges. *Pharmacoeconomics*. 2022;40(6):577–586.
24. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–764.
25. Hernán MA, Robins JM. *Causal inference: what if*. Boco Raton, FL: Chapman & Hall/CRC; 2020.
26. Carpenter JR, Smuk M. Missing data: a statistical framework for practice. *Biom J*. 2021;63(5):915–947.
27. Dahabreh IJ, Robertson SE, Steingrimsson JA, Stuart EA, Hernán MA. Extending inferences from a randomized trial to a new target population. *Stat Med*. 2020;39(14):1999–2014.
28. Degtiar I, Rose S. A review of generalizability and transportability. *Annu Rev Stat Its Appl*. 2021;10(1):501–524.
29. Leahy TP, Kent S, Sammon C, et al. Unmeasured confounding in non-randomized studies: quantitative bias analysis in health technology assessment. *J Comp Eff Res*. 2022;11(12):851–859.
30. Leurent B, Gomes M, Faria R, Morris S, Grieve R, Carpenter JR. Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: a tutorial. *Pharmacoeconomics*. 2018;36(8):889–901.
31. Nguyen TQ, Ackerman B, Schmid I, Cole SR, Stuart EA. Sensitivity analyses for effect modifiers not observed in the target population when generalizing treatment effects from a randomized controlled trial: assumptions, models, effect scales, data scenarios, and implementation details. *PLoS One*. 2018;13(12):e0208795.
32. Dahabreh IJ, Robins JM, Haneuse SJA, et al. Sensitivity analysis using bias functions for studies extending inferences from a randomized trial to a target population. *Stat Med*. 2023;42(13):2029–2043.
33. Wang SV, Pottegård A, Crown W, et al. 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. *Value Health*. 2022;25(10):1663–1672.
34. Hansford HJ, Cashin AG, Jones MD, et al. Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ Open*. 2023;13(9):e074626.
35. NICE real-world evidence framework. Appendix 1 – Data Suitability Assessment Tool (DataSAT). National Institute for Health and Care Excellence. https://www.nice.org.uk/corporate/ecd9/chapter/appendix-1-data-suitability-assessment-tool-datasat. Accessed March 11, 2023.
36. Jaksa A, Wu J, Jónsson P, Eichler HG, Vititoe S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J Comp Eff Res*. 2021;10(9):711–731.
37. *Institute for quality and efficiency in health care. Supply-related data in manufacturer dossiers: things are not yet running smoothly*. Inst Qual Effic Health Care; 2022.
38. NICE real-world evidence framework. National Institute for Health and Care Excellence. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837. Accessed November 2023.
39. Faria R, Hernandex Alava M, Manca A, et al. NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. National Institute for Health and Care Excellence. https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD17-DSU-Observational-data-FINAL.pdf. Accessed November 1, 2022.
40. Phillippo D, Ades T, Dias S, et al. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. National Institute for Health and Care Excellence. https://research-information.bris.ac.uk/ws/portalfiles/portal/94868463/Population_adjustment_TSD_FINAL.pdf. Accessed June 5, 2022.
41. Kuehne F, Arvandi M, Hess LM, et al. Causal analyses with target trial emulation for real-world evidence removed large self-inflicted biases: systematic bias assessment of ovarian cancer treatment effectiveness. *J Clin Epidemiol*. 2022;152:269–280.
42. Amivantamab for treating EGFR exon 20 insertion mutation-positive advanced non-small-cell lung cancer after platinum-based chemotherapy. National Institute for Health and Care Excellence. https://www.nice.org.uk/guidance/ta850/resources/amivantamab-for-treating-egfr-exon-20-insertion-mutationpositive-advanced-nonsmallcell-lung-cancer-after-platinumbased-chemotherapy-pdf-82613545501381. Accessed January , 2023.
43. Mobocertinib for treating EGFR exon 20 insertion mutation-positive advanced non-small-cell lung cancer after platinum-based chemotherapy. National Institute for Health and Care Excellence. https://www.nice.org.uk/guidance/ta855/resources/mobocertinib-for-treating-egfr-exon-20-insertion-mutationpositive-advanced-nonsmallcell-lung-cancer-after-platinumbased-chemotherapy-pdf-82613553899461. Accessed January 16, 2023.
44. FDA granted accelerated approval to Blinatumomab (Blincyto, Amgen Inc.) for the treatment of adult and pediatric patients with B-cell precursor acute lymphoblastic leukemia. US Food & Drug Administration. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-granted-accelerated-approval-blinatumomab-blincyto-amgen-inc-treatment-adult-and-pediatric. Accessed January 5, 2023.
45. Avelumab (BAVENCIO). US Food & Drug Administration. https://www.fda.gov/drugs/resources-information-approved-drugs/avelumab-bavencio. Accessed January 5, 2023.
46. FDA approves alpelisib for PIK3CA-related overgrowth spectrum. US Food & Drug Administration. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-alpelisib-pik3ca-related-overgrowth-spectrum. Accessed January 5, 2023.
47. Vertosick EA, Assel M, Vickers AJ. A systematic review of instrumental variable analyses using geographic region as an instrument. *Cancer Epidemiol*. 2017;51:49–55.
48. Emilsson L, García-Albéniz X, Logan RW, Caniglia EC, Kalager M, Hernán MA. Examining bias in studies of statin treatment and survival in patients with cancer. *JAMA Oncol*. 2018;4(1):63–70.
49. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25(10):1601–1606.
50. Boyne DJ, Cheung WY, Hilsden RJ, et al. Association of a shortened duration of adjuvant chemotherapy with overall survival among individuals with Stage III colon cancer. *JAMA Netw Open*. 2021;4(3):e213587.
51. Admon AJ, Donnelly JP, Casey JD, et al. Emulating a novel clinical trial using existing observational data. Predicting results of the PreVent study. *Ann Am Thorac Soc*. 2019;16(8):998–1007.
52. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology*. 2008;19(6):766–779.
53. Lodi S, Phillips A, Lundgren J, et al. Effect estimates in randomized trials and observational studies: comparing apples with apples. *Am J Epidemiol*. 2019;188(8):1569–1577.
54. Matthews AA, Szummer K, Dahabreh IJ, et al. Comparing effect estimates in randomized trials and observational studies from the same population: an application to percutaneous coronary intervention. *J Am Heart Assoc*. 2021;10(11):e020357.
55. Franklin JM, Patorno E, Desai RJ, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. *Circulation*. 2021;143(10):1002–1013.
56. Wing K, Williamson E, Carpenter JR, et al. Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results. *Health Technol Assess*. 2021;25:1–70.

57. Gatta G, Van Der Zwan JM, Casali PG, et al. Rare cancers are not so rare: the rare cancer burden in Europe. *Eur J Cancer*. 2011;47(17):2493–2511.
58. January CT, Wann LS, Calkins H, et al. AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol*. 2019;74(1):104–132.
59. Clausen M, Mighton C, Kiflen R, et al. Use of real-world evidence in cancer drug funding decisions in Canada: a qualitative study of stakeholders' perspectives. *CMAJ Open*. 2020;8:E772–E778.
60. NICE Decision Support Unit. Technical Support Document. https://nicedsu.sites.sheffield.ac.uk/tsdsr. Accessed May 23, 2023.
61. NICE. Collaboration. *arrangement between: National Institute for Health and Care Excellence, Canadian Agency for Drugs and Technologies in Health, Australian government Department of Health and Aged Care, Healthcare Improvement Scotland, health technology wales (Velindre university NHS Trust), All Wales Therapeutics & Toxicology Centre*. Institut National d'Excellence en Santé et En Services Sociaux, Pharmac. 2023.
62. Regulation on health technology assessment. European Commission. https://health.ec.europa.eu/health-technology-assessment/regulation-health-technology-assessment_en#:~:text=The%20Regulation%20(EU)%202021%2F,of%20HTA%20across%20the%20Union. Accessed July 31, 2023.
63. EUnetHTA. D4.3 Direct and Indirect Comparisons. https://www.eunethta.eu/d4-3/. Accessed February 9, 2023.
64. *EUnetHTA. Vision Paper on the Sustainable Availability of the Proposed Registry Evaluation and Quality Standards Tool (REQueST)*. 2019.
65. Padula WV, Kreif N, Vanness DJ, et al. Machine learning methods in health economics and outcomes research-the PALISADE checklist: a good practices report of an ISPOR task force. *Value Health*. 2022;25(7):1063–1080.
66. Birnbaum B, Nussbaum N, Seidl-Rathkopf K, et al. Model-assisted cohort selection with bias analysis for generating large-scale cohorts from the EHR for oncology research. arXiv. Preprint. https://arxiv.org/abs/2001. Accessed February 10, 2023.
67. DATA-CAN, DATA-CAN: The Health Data Research Hub for Cancer. https://www.data-can.org.uk/t. Accessed October 7, 2022.
68. NHS Digital. Secure Data Environment service. https://digital.nhs.uk/services/secure-data-environment-servicet. Accessed July 24, 2023.
69. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ*. 2021;373:n826.
70. EHDEN. EHDEN Database Catalogue. https://www.ehden.eu/ehden-portal/t. Accessed October 7, 2022.
71. Kent S, Burn E, Dawoud D, et al. Common problems, common data model solutions: evidence generation for health technology assessment. *Pharmacoeconomics*. 2021;39(3):275–285.
72. Observational Health Data Sciences and Informatics (OHDSI) initiative. OMOP Common Data Model. https://www.ohdsi.org/data-standardization/the-common-data-model/t. Accessed October 7, 2022.
73. European Medicines Agency, Heads of Medicines Agencies. Data Quality Framework for EU Medicines Regulation. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation_en.pdfn. Accessed March 2, 2023.
74. Leahy TP, Duffield S, Kent S, et al. Application of quantitative bias analysis for unmeasured confounding in cost-effectiveness modelling. *J Comp Eff Res*. 2022;11(12):861–870.
75. Popat S, Liu SV, Scheuer N, et al. Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of Pralsetinib in non-small cell lung cancer. *Nat Commun*. 2022;13(1):3500.
76. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using real-world data to predict findings of an ongoing Phase IV cardiovascular outcome trial: cardiovascular safety of linagliptin versus glimepiride. *Diabetes Care*. 2019;42(12):2204–2210.
77. RCT Duplicate. *Predicting Findings of Ongoing Phase 4 RCTs With Real World Data Analyses: an Assessment to Support FDA's Regulatory Decision Making*. 2023.
78. RCT Duplicate. Other Prospective Replications. https://www.rctduplicate.org/other-prospective-projects.htmlt. Accessed July 24, 2023.
79. Crown W, Dahabreh IJ, Li X, Toh S, Bierer B. Can observational analyses of routinely collected data emulate randomized trials? Design and feasibility of the observational patient evidence for regulatory approval science and understanding disease project. *Value Health*. 2023;26(2):176–184.
80. Fu EL, Evans M, Carrero JJ, et al. Timing of dialysis initiation to reduce mortality and cardiovascular events in advanced chronic kidney disease: nationwide cohort study. *BMJ*. 2021;375:e066306.
81. Josey KP, Yang F, Ghosh D, Raghavan S. A calibration approach to transportability and data-fusion with observational data. *Stat Med*. 2022;41(23):4511–4531.
82. Ali MS, Prieto-Alhambra D, Lopes LC, et al. Propensity score methods in health technology assessment: principles, extended applications, and recent advances. *Front Pharmacol*. 2019;10:973.
83. Chernozhukov V, Chetverikov D, Demirer M, et al. *Double/Debiased Machine Learning for Treatment and Structural Parameters*. Oxford, United Kingdom: Oxford University Press; 2018.
84. Moler-Zapata S, Grieve R, Lugo-Palacios D, et al. Local instrumental variable methods to address confounding and heterogeneity when using electronic health records: an application to emergency surgery. *Med Decis Making*. 2022;42(8):1010–1026.
85. EHDEN. EHDEN Academy. https://academy.ehden.eu/. Accessed December 1, 2022.
86. Innovative Health Initiative. *The GetReal Initiative*. 2018.
87. European Commission. *New Methods for the Effective Use of Real-World Data and/or Synthetic Data in Regulatory Decision-Making and/or in Health Technology Assessment*. 2023.
88. FDA. RWD and RWE-Focused Demonstration Projects. https://www.fda.gov/science-research/real-world-evidence/rwd-and-rwe-focused-demonstration-projects. Accessed July 31, 2023.
89. NICE. Our Projects and Partners. https://www.nice.org.uk/about/what-we-do/our-research-work/our-projects-and-partners. Accessed July 31, 2023.