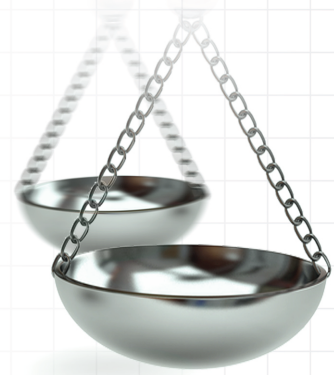
White Paper

For reprint orders, please contact: reprints@futuremedicine.com



Journal of **Comparative Effectiveness Research**

The use of nonrandomized evidence to estimate treatment effects in health technology assessment

Seamus Kent*[,1], Maximilian Salcher-Konrad[2,3], Stefania Boccia[4,5], Jacoline C Bouvy[1], Chiara de Waure[6], Jaime Espin[7,8,9], Karen Facey[10], Mary Nguyen[3], Juan Carlos Rejon-Parrilla[11] & Pall Jonsson[1]

[1]National Institute for Health & Care Excellence, Manchester, M1 4BT, UK
[2]Care Policy & Evaluation Center (CPEC), London School of Economics & Political Science, London, WC2A 2AE, UK
[3]LSE Health, London School of Economics & Political Science, London, WC2A 2AE, UK
[4]Section of Hygiene, University Department of Life Sciences & Public Health, Università Cattolica del Sacro Cuore, Rome, 20123, Italy
[5]Department of Woman & Child Health & Public Health – Public Health Area, Fondazione Policlinico Universitario A Gemelli IRCCS, Rome RM, 00168, Italy
[6]Department of Medicine & Surgery, University of Perugia, Perugia, 06123, Italy
[7]Andalusian School of Public Health/Escuela Andaluza de Salud Pública (EASP), Granada, 18011, Spain
[8]CIBER en Epidemiología y Salud Pública (CIBERESP), Spain / CIBER of Epidemiology & Public Health (CIBERESP), Planta 0, Madrid, 28029 Spain
[9]Instituto de Investigación Biosanitaria ibs, Granada, 18012, Spain
[10]Usher Institute, University of Edinburgh, Nine Bioquarter, 9 Little France Road, Edinburgh, EH16 4UX, UK
[11]Andalusian HealthTechnology Assessment Area, Fundación Progreso y Salud (AETSA-FPS), Sevilla,41092, Spain
*Author for correspondence: Seamus.Kent@nice.org.uk

Health technology assessment (HTA) is increasingly informed by nonrandomized studies, but there is limited guidance from HTA bodies on expectations around evidence quality and study conduct. We developed recommendations to support the appropriate use of such evidence based on a pragmatic literature review and a workshop involving 16 experts from eight countries as part of the EU's Horizon-2020 IMPACT-HTA program (work package six). To ensure HTA processes remain rigorous and robust, HTA bodies should demand clear, extensive and structured reporting of nonrandomized studies, including an in-depth assessment of the risk of bias. In recognition of the additional uncertainty imparted by nonrandomized designs in estimates of treatment effects, HTA bodies should strengthen early scientific advice and engage in collaborative efforts to improve use of real-world data.

First draft submitted: 4 May 2021; Accepted for publication: 28 June 2021; Published online: 19 July 2021

Keywords: comparative effectiveness research • health technology assessment • nonrandomized trials • observational research • real-world evidence

Most health technology assessment (HTA) bodies profess a strong preference for evidence on treatment effects from randomized controlled trials (RCTs) over nonrandomized studies due to a lower risk of bias by design [1]. However, with an increasing number of medicines receiving regulatory approval based on nonrandomized evidence, HTA bodies are being tasked with issuing recommendations on new technologies in the absence of RCTs or with limited RCT data [2–5]. For other health technologies like medical devices, nonrandomized studies already provide the predominant source of evidence [3]. Nonrandomized studies are defined broadly here to include all study designs without randomization, including nonrandomized clinical trials, observational studies and trials with external controls [4].

The quality of nonrandomized studies so far accepted by HTA bodies for decision making has been variable [2–5,5] and there is a widely acknowledged lack of transparency in research governance throughout the evidence generation process [6,7]. In this paper we present recommendations aimed both at those generating evidence and HTA bodies

Future Medicine

to ensure that where nonrandomized studies are of potential value, they are conducted in transparent and robust ways to maximize their usefulness for decision making.

The use of nonrandomized evidence to estimate treatment effects

RCTs are widely recognized as the gold standard for estimating treatment efficacy because of the ability of randomization in ensuring any differences in baseline characteristics between groups are due to chance, blinding (where applied) in preventing knowledge of treatment allocation from influencing behaviors and standardized protocols in ensuring consistent data collection [8]. RCTs are not; however, immune to bias. More than half of the pivotal studies in oncology submitted to the EMA between 2014 and 2016 were judged to be at high risk of bias for their primary end point [9]. Further, many well-conducted RCTs have limitations with regard to establishing comparative effectiveness for HTA including comparisons to treatments (or placebo) other than usual care, selected patient populations, insufficient follow-up, treatment protocols that deviate from usual care and the use of surrogate outcomes which may not predict the outcome(s) of ultimate interest. Against this background, there is at least the potential that well-conducted nonrandomized studies based on high quality and relevant (e.g., inclusion of patient relevant outcomes) data could improve decision making.

In addition, RCTs may be considered unethical (e.g., last line cancer therapies) or infeasible (e.g., due to small numbers of eligible patients) [10]. There has been an increase in the use of expedited access programs by regulatory authorities to facilitate faster patient access to innovative treatments, particularly in oncology and rare diseases, where single-arm trials are common [11]. Following approval through expedited regulatory pathways, HTA bodies are required to make initial reimbursement and/or pricing decisions based on this limited evidence [12–15]. When substantial uncertainties exist about the clinical evidence for a highly promising treatment, and data collection to resolve the uncertainties is feasible, a conditional reimbursement decision may be given, requiring an outcomes-based managed entry agreement to collect data post reimbursement to resolve the uncertainties, commonly in the form of a nonrandomized study [13].

Nonrandomized studies are typically at a higher risk of bias than RCTs because patients are not randomly allocated to treatment and treatment allocation is known; instead, physicians decide on a patient's treatment based on their expectation of the benefit–risk profile of different treatments for that patient and their preferences [14]. Bias may also arise from poor data quality including errors in data entry, measurement error, misclassification of exposures and outcomes, missing data or from poor analytical choices, coding errors or selective reporting and/or publication of results [15].

There is mixed evidence on the internal validity of nonrandomized studies. A 2017 review of 14 meta-epidemiological studies reported that seven found no systematic differences between RCTs and nonrandomized studies, five found estimates from nonrandomized studies to systematically exceed those from RCTs and the remaining two were inconclusive [16]. A recent meta-epidemiological study conducted by IMPACT-HTA work package six consortium found no evidence of a systematic difference in treatment effects for pharmaceutical interventions but substantial variation in these differences for specific clinical questions. Similar results were observed among the first ten trial emulations from the RCT-DUPLICATE project in which RCTs were replicated in US claims databases [17]. There is stronger evidence that nonrandomized studies relying on external controls, such as single-arm trials, are associated with greater bias on average than other types of nonrandomized studies [16].

Despite the increasing availability of observational data and advances in analytical methods, RCTs remain the main source of evidence on treatment effects for medicines [18]. Where nonrandomized studies are used, they are usually based on single-arm trials and have predominantly been in oncology, infectious diseases and for rare disease treatments. The proportion of submissions using nonrandomized evidence is increasing: more than half of all NICE's medicines appraisals between 2010 and 2016 using nonrandomized data on clinical effectiveness were in 2015–16 alone [19]. For other types of health technologies like medical devices, there has long been a greater reliance on nonrandomized studies to estimate treatment effects as regulatory requirements do not generally demand RCTs and RCTs may be difficult to perform [3].

Despite the pervasive risk of bias from confounding in nonrandomized studies, many HTA bodies and regulators have accepted submissions using no or only simple methods of adjustment [5,19]. Anderson *et al.* [19] found that of the 22 medicines appraised by NICE using nonrandomized data between 2010 and 2016, only five used a regression approach to adjust for confounding while two-thirds used naive unadjusted indirect comparison to aggregate data. A persistent challenge is that for clinical questions where RCTs cannot be performed due to small patient numbers, there is also likely to be limited observational data, inhibiting the extent of adjustment possible in analyses.

Recommendations for the use of nonrandomized studies in comparative effectiveness estimation

Materials & methods

The work consisted of two key activities. First, we undertook two pragmatic reviews. The first focused on empirical assessments of the internal validity of treatment effect estimates from nonrandomized studies. We reviewed such studies that were included in a Cochrane umbrella review published in 2014 [20] and another, pragmatic review published in 2017 [21]. To capture more recently published studies, we also replicated the search from the Cochrane review for one database (MEDLINE via PubMed) up until September 2018. Our second review focused on best-practice recommendations for the generation of evidence on the effects of new technologies from reports and guidance from HTA bodies and societies, payers, regulators and real-world evidence research consortia including CADTH, European Network for Health Technology Assessment (EUnetHTA), Haute Autorite de Sante (HAS), Health Information and Quality Authority (HIQA), Institute for Clinical and Economic Review (ICER), Rijksinstituut voor Ziekte en Invaliditeitsverzekering (INAMI-RIZIV), Institut fur Qualitat und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), National Institute for Health and Care Excellence (NICE), Pharmaceutical Benefits Advisory Committee (PBAC), Scottish Medicines Consortium (SMC), Tandvards-och lakemedelsformansverket (TLV), Zorginstituut Nederlands (ZIN), US FDA, EMA, Duke Margolis, ENCePP, GetReal, HTAi, ISPE, ISPOR, RWE4Decisions and REALISE. We identified additional papers and reports through the reference lists of identified sources and through discussion with experts. We also conducted *ad-hoc* literature reviews to identify additional academic papers on specific topics. The review was restricted to English language reports, which represents a limitation of this work.

We drafted an initial list of key recommendations based on our literature review. These were revised based on consultation with European HTA bodies and regulators. We held an online workshop in June 2020 involving 16 participants (see acknowledgements) from eight countries across Europe (the UK, the Netherlands, Spain, Germany, Norway, Sweden, Austria and Italy). Initial recommendations were shared with all participants prior to the meeting and the revised set were sent for written comment after the meeting. Written comments were incorporated to arrive at the list of considerations presented in this paper.

Recommendations

Conduct of nonrandomized studies

Planning & design

Justify the need for a nonrandomized study & demonstrate that the research question is amenable to being answered using nonrandomized data.

The need for a nonrandomized study should be justified and research questions should be clearly defined using established frameworks [22,23]. Where observational data are considered for use as part of a nonrandomized study, it is important to ascertain that: the question is amenable to being answered using observational data [24], and that one or more data sources containing sufficient information of high quality and relevant to the decision context are available and accessible [25]. Datasets should be identified through a systematic, transparent and reproducible process to ensure the most appropriate data are used [26,27]. This avoids the selection of datasets based on convenience or the knowledge or expectation of deriving particular results.

Plan studies prospectively and engage in early scientific advice procedures

Nonrandomized studies should be planned prospectively to negate the possibility of selective methodology and results [7]. Study sponsors should take advantage of scientific consultation processes where available to help guide evidence generation throughout a product's lifecycle. Scientific advice provides an essential function in allowing HTA bodies and regulators to work with stakeholders to ensure that plans for evidence generation deliver information that supports and improves decision making in both premarket and postmarket settings [28,29].

Analysis

Understand potential risks of bias & address using appropriate analytical strategies

Bias may arise in nonrandomized studies for numerous reasons including patient selection, confounding or data limitations [30]. The potential mechanisms of bias for any application should be clearly articulated and the analysis plan designed to elucidate and minimize potential bias. Analysts can seek to address bias through the application of statistical methods or study design. An increasingly common approach is to try to replicate RCT designs – the

'target trial' approach [31]. This involves numerous design principles, but key among these are restricting cohorts to new users, using active comparators (as opposed to nonusers), adjusting based on pretreatment confounders (e.g., using propensity score matching), performing or utilizing outcome validation studies and prespecifying sensitivity analyses. In practice, this can be difficult to achieve given the limitations of many observational datasets. In addition, for novel therapies and for certain conditions, it may not be possible to identify appropriate active comparators, while early users of a technology may differ from the target population [32]. Where using external control arms, high-quality individual patient level data (rather than aggregate data) from contemporaneous (or where necessary, historical) controls should be used for adjustment as it provides greater scope to control for differences between patients [33]. Regardless of study design, potential confounders should typically be defined prior to study conduct and based on the scientific literature and through engagement with clinical experts [34].

Perform extensive sensitivity analyses

Nonrandomized studies involve many decisions and assumptions including in data curation and analysis, each of which, alone or in combination, could have substantial effects on the resulting estimates. It is therefore, essential that extensive sensitivity analyses are undertaken to understand the robustness of the results to data curation, design and analysis decisions and characterize the uncertainty in the treatment effect [35,36]. This could be complemented by quantitative bias analysis, which includes techniques such as negative controls, the use of external information and threshold analysis, in other words, identifying the extent of bias sufficient to change decisions [37]. Finally, where possible there should be an attempt to replicate findings in one or more datasets [38].

Reporting

Register protocols before study conduct

Detailed study protocols including statistical analysis plans should be registered before the beginning of the study on publicly accessible platforms using structured reporting templates, and amendments to the protocols should be clearly reported and justified [7]. This would improve study transparency and allay concerns about selective analyses and selective reporting as well as publication bias, which are major impediments to the acceptability and wider use of nonrandomized evidence [39].

Report data, methods & results transparently

It is essential that all information pertaining to evidence quality is clearly and comprehensively reported. This includes descriptions of data fitness-for-purpose, encompassing data quality and relevance, reporting of study conduct including data curation, analysis and results. Traditional reporting checklists play an important role in the transparent reporting of nonrandomized study methods and results by ensuring that key information is reported [40]. While they should be used in evidence submissions, they are not generally sufficient to support reproduction and are not an indicator of quality. Novel approaches to transparent reporting provide structured templates to articulate study design and analysis and the assumptions underlying them [41]. Ideally reviewers of submitted evidence, including HTA bodies or independent review groups, would also have access to the data and analytical code to ensure the replicability of the submitted results and assess the impact of alternative analytical decisions or data on the resulting estimate(s). However, there remain substantial governance, technical and practical challenges to sharing data, including a lack of in-house expertise in many HTA agencies [42]. Validated analytical platforms can complement preregistration in promoting transparency by ensuring a correct ordering and comprehensive account of study conduct, while reducing the risk of analytical or coding errors [35].

Describe potential biases & report the overall risk of bias

Study sponsors should clearly articulate potential causes of bias and their impact on estimated treatment effects. The overall risk of bias should also be formally assessed using well-validated tools [43]. For nonrandomized studies, the ROBINS-I tool is recommended by the European network for HTA (EUnetHTA) and assesses the risk of bias by specifying the research question as a target trial and considering risks from seven domains of bias namely bias due to confounding, in selection of participants into the study, in classification of interventions, due to deviations from intended interventions, due to missing data, in measurement of outcomes and in the selection of the reported result [44]. Other sources of bias beyond those covered within existing quality assurance tools should also be documented [43]. Ideally, quantitative estimates of bias would also be presented [37].

Convey & ideally quantify the uncertainty

Uncertainty is pervasive in HTA, particularly where estimates of effects and costs are required over the long-term. Although several meta-epidemiological studies, as well as trial replication studies, have found no evidence of systematic differences in treatment effect estimates between RCTs and nonrandomized studies, there was great variation in estimates across clinical questions. Unless a decision maker is able to identify when differences are likely to be present and the direction and magnitude of any differences, then they must acknowledge the sizeable additional uncertainty involved in the use of nonrandomized evidence which will not be fully captured by the statistical uncertainty in the estimated effectiveness parameter [45]. This uncertainty should be appropriately conveyed and ideally quantified, in evidence submissions.

Further considerations for HTA bodies

Strengthening systems

Strengthen & standardize scientific advice procedures

Scientific advice provides an opportunity to enhance quality and relevance of evidence submitted to HTA bodies, but they are often limited to discussion of RCTs. HTA bodies should consider extending their use to discuss any nonrandomized evidence that will be important in the determination of value and ensure that experts providing the advice have expertise in the evaluation of such studies. Furthermore, international scientific advice processes are to be encouraged with collaboration over best practice guidelines that can be referred to in relation to the design and conduct of nonrandomized studies, such as those presented in this paper.

Strengthen conditional reimbursement processes to ensure generation of further informative evidence after initial reimbursement decisions

Uncertainty poses the risk that incorrect decisions will be made that are detrimental to population health. Managed entry agreements of various forms have the potential to ameliorate the impact of uncertainty and are widely used throughout Europe [28,46]. A perceived lack of transparency around such schemes means there has been an incomplete understanding of their strengths and limitations. Some outcomes-based managed entry agreements that have sought to collect data for later reappraisal have been compromised by several procedural and methodological limitations with poor quality studies commissioned and limited impact on decision making [47]. These experiences underscore the importance of having high-quality evidence available at the time of the initial approval. When such outcomes-based managed entry agreements are used, it is imperative that clear responsibilities for data collection and analysis are defined and that appropriate enforcement mechanisms are available to HTA bodies to ensure timely delivery of high-quality evidence. Data collection and analysis should follow best practice guidance and be reported and published transparently. Further guidance is provided in work package ten of the IMPACT HTA project [48].

Invest in & develop staff skills in the design, analysis & interpretation of nonrandomized studies

As an increasing number of HTA submissions are made on the basis of nonrandomized data and as HTA bodies respond by commissioning and designing studies and perhaps even analyzing data, it is imperative that their staff and decision making committees possess the requisite skills to conduct this work [39]. This is likely to require both training and recruitment. International collaboration with experts is also essential to share expertise and develop methods specifically for HTA.

Issuing & enforcing best practice guidance

To ensure the generation of high-quality evidence suitable for decision making, HTA bodies should issue clear guidance on data quality standards and best practice methods for the design, conduct and reporting of nonrandomized studies and ensure that these are followed. There are several ongoing initiatives to establish frameworks for the use of real-world evidence in decision making including by the FDA, EMA and NICE [13,49,50]. It is essential that these tools are built in a collaborative fashion to streamline evidence generation and ensure adoption. They should aspire to be simple to apply and interpret so as to enhance transparency and reduce the burden on sponsors and decision makers. This would also benefit smaller HTA bodies and payers who have fewer resources to conduct detailed assessments themselves. This work could also include clear guidance as to when nonrandomized studies will be considered and where they are expected. Because this is an evolving field, it is important that these frameworks are regularly revised.

Supporting future research & initiatives

Support access to high-quality data

HTA bodies should support international efforts to set data standards; thereby, improving the quality and interoperability of data. They should also support initiatives to improve data linkage and access while ensuring data protection is maintained which could include the exploration of trusted research environments, federated data networks and the development of high-quality synthetic datasets.

Support methodological and empirical research

HTA bodies should further support research efforts to advance methods for causal estimation and evidence synthesis and empirical investigations into the performance of different nonrandomized study designs across different use cases to better understand in which circumstances nonrandomized studies are most likely to give reliable answers and which design and analytical features are most important for generating robust estimates of treatment effects.

Conclusion

While most HTA bodies have a strong preference for evidence on treatment effects to be derived from RCTs, they are increasingly being asked to make reimbursement and/or pricing decisions based on nonrandomized studies. These studies are at higher risk of bias than RCTs meaning estimates of clinical effectiveness are often highly uncertain. When RCT evidence is insufficient for decision making, rigorous and extensive processes should be followed to ensure that evidence derived from nonrandomized studies is of high quality and those conducting such studies adhere to best practices, including the use of high-quality data, addressing bias and confounding using appropriate methods and transparency in study design, conduct, analysis and reporting. Even with high-quality research, HTA bodies should recognize the uncertainty inherent in nonrandomized studies and establish robust mechanisms to mitigate the risks for population health thereby imposed.

Executive summary

- There is increasing interest in estimating treatment effects from nonrandomized studies for health technology assessment (HTA), but limited guidance from HTA bodies on expectations around evidence quality and study conduct.
- We developed several recommendations aimed at both evidence developers and HTA bodies to improve the quality and value of nonrandomized studies.
- These recommendations were informed by a pragmatic literature review and a workshop involving 16 HTA experts from eight European countries.
- To ensure HTA processes remain rigorous and robust, HTA bodies should demand clear, extensive and structured reporting of nonrandomized studies, including an in-depth assessment of the risk of bias.
- In recognition of the additional uncertainty imparted by nonrandomized designs in estimates of treatment effects, HTA bodies should strengthen early scientific advice and consider using managed entry agreements that may help mitigate clinical uncertainty if well designed.
- HTA bodies must also ensure that staff are equipped with the requisite skills to critically appraise nonrandomized studies.

Author contributions

S Kent and M Salcher-Konrad drafted the manuscript. All other co-authors commented on multiple versions of the manuscript. All co-authors participated in the workshop to discuss the draft recommendations.

Acknowledgments

The authors would like to thank all attendees at the June 2020 workshop who provided an excellent and thorough discussion on this topic. Attendees were A Jonas, A Schiel, B Wieseler, C de Waure, D Lundin, J Bouvy, J Espin, JC Rejon, K Facey, M Nguyen, M Salcher-Konrad, M Walter, N Hedberg, P Jonsson, S Boccia and W Goettsch. Their contributions to the workshop should not be interpreted as an endorsement of the manuscript.

Financial & competing interests disclosure

This project has received funding from the EU's Horizon 2020 research and innovation program under grant agreement no: 779312 (www.impact-hta.eu). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Makady A, ten Ham R, 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* 20, 520–532 (2017).
2. Hatswell AJ, Baio G, Berlin JA *et al.* Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999–2014. *BMJ Open* 6, e011666 (2016).
3. Crispi F, Naci H, Barkauskaite E *et al.* Assessment of devices, diagnostics and digital technologies: a review of NICE Medical Technologies Guidance. *Appl. Health Econ. Health Pol.* 17, 189–211 (2019).
4. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet* 359, 57–61 (2002).
5. Hatswell AJ, Freemantle N, Baio G. Economic evaluations of pharmaceuticals granted a marketing authorisation without the results of randomised trials: a systematic review and taxonomy. *Pharmacoeconomics* 35, 163–176 (2017).
6. ISPOR. Improving transparency in non-interventional research for hypothesis testing – why, what, and how: considerations from the real-world evidence transparency initiative (draft White Paper) (2019). www.ispor.org/docs/default-source/strategic-initiatives/improving-transparency-in-non-interventional-research-for-hypothesis-testing_final.pdf?sfvrsn=77fb4e97_6
7. Orsini LS, Berger M, Crown W *et al.* 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* 23, 1128–1136 (2020).
8. Collins R, Bowman L, Landray M *et al.* The magic of randomization versus the myth of real-world evidence. *N. Engl. J. Med.* 382, 674–678 (2020).
9. Naci H, Davis C, Savović J *et al.* Design characteristics, risk of bias, and reporting of randomised controlled trials supporting approvals of cancer drugs by European Medicines Agency, 2014–16: cross sectional analysis. *BMJ* 366, l5221 (2019).
10. Eichler HG, Koenig F, Arlett P *et al.* Are novel, nonrandomized analytic methods fit for decision making? The need for prospective, controlled, and transparent validation. *Clin. Pharmacol. Ther.* 107, 773–779 (2019).
11. 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* 9, e024895 (2019).
12. Nicod E, Annemans L, Bucsics A *et al.* HTA programme response to the challenges of dealing with orphan medicinal products: process evaluation in selected European countries. *Health Pol.* 123, 140–151 (2019).
13. Facey K, Rannanheimo P, Batchelor L *et al.* Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU – actions for stakeholders. *Int. J. Technol. Assess. Health Care* 36, 459–468 (2020).
14. Gerstein HC, McMurray J, Holman RR. Real-world studies no substitute for RCTs in establishing efficacy. *Lancet* 393, 210–211 (2019).
15. Asche CV, Seal B, Kahler KH *et al.* Evaluation of healthcare interventions and big data: review of associated data issues. *Pharmacoeconomics* 35, 759–765 (2017).
16. Woolacott N, Corbett M, Jones-Diette J *et al.* Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. *J. Clin. Epidemiol.* 90, 108–118 (2017).
17. 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* 143(10), 1002–1013 (2021).
18. 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* 36, 359–368 (2018).
19. Anderson M, Naci H, Morrison D *et al.* A review of NICE appraisals of pharmaceuticals 2000–2016 found variation in establishing comparative clinical effectiveness. *J. Clin. Epidemiol.* 105, 50–59 (2019).
20. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst. Rev.* 4, MR000034 (2014).
21. Woolacott N, Corbett M, Jones-Diette J *et al.* Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. *J. Clin. Epidemiol.* 90, 108–118 (2017).

22. ICH. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials E9(R1). *International Conference on Harmonization* (2019). https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf
23. Agency for Healthcare Research and Quality. Developing a protocol for observational comparative effectiveness research: a user's guide. (2013). https://effectivehealthcare.ahrq.gov/products/observational-cer-protocol/research
24. Bartlett VL, Dhruva SS, Shah ND *et al.* Feasibility of using real-world data to replicate clinical trial evidence. *JAMA Netw. Open* 2, e1912869 (2019).
25. Franklin JM, Glynn RJ, Martin D *et al.* Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. *Clin. Pharmacol. Ther.* 105, 867–877 (2019).
26. Eichler HG, Bloechl-Daum B, Bauer P *et al.* "Threshold-crossing": a useful way to establish the counterfactual in clinical trials? *Clin. Pharmacol. Ther.* 100, 699–712 (2016).
27. Gatto NM, Reynolds RF, Campbell UB. A structured preapproval and post approval comparative study design framework to generate valid and transparent real-world evidence for regulatory decisions. *Clin. Pharmacol. Ther.* 106, 103–115 (2019).
28. Moseley J, Vamvakas S, Berntgen M *et al.* Regulatory and health technology assessment advice on postlicensing and postlaunch evidence generation is a foundation for lifecycle data collection for medicines. *Br. J. Pharmacol.* 86, 1034–1051 (2020).
29. Ng T, Ziomek J, Delaitre-Bonnin C *et al.* PHP276 – a review of the impact of integrated scientific advice for the optimisation of evidence generation for HTA appraisals. *Value Health* 21, S197 (2018).
30. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). The European Network of Centres for pharmacoepidemiology and pharmacovigilance (ENCePP) guide on methodological standards in pharmacoepidemiology (revision 8) (2020). www.encepp.eu/standards_and_guidances/documents/GuideMethodRev8.pdf
31. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidemiol.* 183, 758–764 (2016).
32. Schneeweiss S, Gagne JJ, Glynn RJ *et al.* Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin. Pharmacol. Ther.* 90, 777–790 (2011).
33. European Medicines Agency. Workshop on single-arm trials in oncology (2016). www.ema.europa.eu/en/events/workshop-single-arm-trials-oncology
34. Institute for Quality and Efficiency in Health Care. [A19-43] Development of scientific concepts for the generation of routine practice data and their analysis for the benefit assessment of drugs according to §35a Social Code Book V – rapid report (2020). www.iqwig.de/en/projects-results/projects/drug-assessment/a19-43-development-of-scientific-concepts-for-the-generation-of-routine-practice-data-and-their-analysis-for-the-benefit-assessment-of-drugs-according-to-35a-social-code-book-v-rapid-report
35. Schneeweiss S. Real-World Evidence of Treatment Effects: The Useful and the Misleading. *Clin. Pharmacol. Ther.* 106, 43–44 (2019).
36. Lawlor DA, Tilling K, Smith GD. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* 45, 1866–1886 (2016).
37. Sammon CJ, Leahy TP, Gsteiger S *et al.* Real-world evidence and nonrandomized data in health technology assessment: use existing methods to address unmeasured confounding? *J. Comp. Eff. Res.* 9(14), 969–972 (2020).
38. Berger ML, Sox H, Willke RJ *et al.* 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. *Value Health* 20, 1003–1008 (2017).
39. HTAi Global Policy Forum. Real-world evidence in the context of health technology assessment processes – from theory to action (2018). https://htai.org/wp-content/uploads/2019/02/HTAiGlobalPolicyForum2019_BackgroundPaper.pdf
40. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). ENCePP Resources Database (2020). www.encepp.eu/encepp/resourcesDatabase.jsp
41. Wang SV, Pinheiro S, Hua W *et al.* STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ* 372, m4856 (2021).
42. Wang SV, Schneeweiss S, Berger ML *et al.* Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Value Health* 20, 1009–1022 (2017).
43. 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* 11, e043961 (2021).
44. Sterne JA, Hernán MA, Reeves BC *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355, i4919 (2016).
45. Deeks JJ, Dinnes J, D'Amico R *et al.* Evaluating non-randomised intervention studies. *Health Technol. Assess.* 7, 1–186 (2003).
46. Kanavos P, Ferrario A, Tafuri G *et al.* Managing risk and uncertainty in health technology introduction: the role of managed entry agreements. *Global Pol.* 8, 84–92 (2017).
47. Wenzl M, Chapman S. Performance-based managed entry agreements for new medicines in OECD countries and EU member states: how they work and possible improvements going forward (2019) (OECD Health Working Papers, No.

115). www.oecd-ilibrary.org/social-issues-migration-health/performance-based-managed-entry-agreements-for-new-medicines-in-oecd-countries-and-eu-member-states_6e5e4c0f-en

48. Facey K, Whittal A, Drummond M *et al.* IMPACT HTA WP10 HTA appraisal framework suitable for rare disease treatments (2021). https://8c3e11d9-5f36-452f-abe3-c95befd6e85d.filesusr.com/ugd/e1a359_70bcc08bb2584a10b1717000e5d14fc9.pdf

49. Head of Medicines Agencies and European Medicines Agency. HMA-EMA Joint Big Data Taskforce Phase II report: 'Evolving Data-Driven Regulation' (2020). www.ema.europa.eu/en/documents/other/hma-ema-joint-big-data-taskforce-phase-ii-report-evolving-data-driven-regulation_en.pdf

50. National Institute for Health and Clinical Excellence. Increasing use of health and social care data in guidance development (2019). www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-guidelines/how-we-develop-nice-guidelines/data-and-analytics-statement-of-intent