Title: Structure and Content of a Taxonomy to Support the Use of Real-World Evidence by Health Technology Assessment Practitioners and Healthcare Decision Makers

Précis: A practical approach of linking questions asked of real-world evidence with specific data sources to provide examples and learnings to support better use of RWE in HTA

Linda A. Murphy, PhD, Lumanity Europe, Steel City House, West Street, Sheffield, England UK. Linda.Murphy@lumanity.com

Ron Akehurst, DSc

Oriol Solà-Morales, PhD

David Cunningham, MD

Jorge Mestre-Ferrandiz, PhD

Matthew Franklin, PhD

Gérard de Pouvourville, PhD

Authorship contributions

Concept and design: Akehurst, Murphy, Solà-Morales, Cunningham, Mestre-Ferrandiz, de Pouvourville

Acquisition of data: Murphy, Franklin

Analysis and interpretation of data: Akehurst, Murphy, Solà-Morales, Cunningham, Mestre-Ferrandiz, Franklin, de Pouvourville

Drafting of the manuscript: Akehurst, Murphy, Solà-Morales, Cunningham, Mestre-Ferrandiz, Franklin, de Pouvourville

Critical revision of the paper for important intellectual content: Akehurst, Murphy, Solà-Morales, Cunningham, Mestre-Ferrandiz, Franklin, de Pouvourville

Obtaining funding: Akehurst, Murphy, Solà-Morales, Cunningham, Mestre-Ferrandiz, de Pouvourville

Administrative, technical, or logistic support:

Conflict of interest statement

Acknowledgements

Title Structure and Content of a Taxonomy to Support the Use of Real-World Evidence by Health Technology Assessment Practitioners and Healthcare Decision Makers

Description of the article

Our work, and that of other researchers, has highlighted the barriers to the optimal use of real-world evidence (RWE) in health technology assessment (HTA). When RWE has been used, particularly to address questions of relative effectiveness in HTA, it has commonly been mistrusted. This is in part because problems nearly always arise that are particular to the use of particular data sources, answering particular questions. Each 'pairing' of question and source has its own problems. We have argued in other papers in this supplement for the potential value of a resource that allows easy retrieval of past decisions by HTA agencies classified by 'pairings'. This paper sets out to describe in detail lacking in the other papers such a practical resource, which we would be happy to make available to the HTA community to own and use.

Abstract

This is the third in a series of papers that consider the barriers to optimal use of real-world evidence (RWE) in health technology assessment (HTA) and how to overcome them. The work was carried out as part of EUreccA 2025, in particular with the RWE workstream embodied within that collaboration. In Paper 1 we described the reasoning and process that led us to develop practical tools to support RWE use, including this taxonomy. The taxonomy classifies questions that are typically addressed using real-world data (RWD) in HTA and the data sources typically used to address these questions. In Paper 2 we explained the methods used to develop the taxonomy and other outputs from this work. In this paper, we describe the taxonomy itself. For as many of the pairings as possible, we have provided links to advice and methods on how to address the associated question using those data. We have also provided links to examples of RWE use in practical decision making to answer the questions posed. Our work is not complete, but we believe it is sufficient to demonstrate the value of such a taxonomy and information source if it is completed and curated as a 'wiki' by the community that would use it.

1. Background

1.1 Broad context of evolving RWE efforts

There have been increasing calls for greater use of real-world data and evidence (RWD/E), to enable faster provision of effective medicines to patients in need. However, RWE, particularly when used to estimate the relative effectiveness of interventions, is not always readily accepted by agencies responsible for reimbursement and pricing of new pharmaceuticals. This issue remains despite an increasing number of initiatives or associations aiming to improve the standards of collection and analysis of RWD, to support confidence in its use.[2-10] These include RWE4Decisions, GetReal, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Transparency Initiative, CanREValue, and the European Federation of Pharmaceutical Industries and Associations (EFPIA), to name a few. Within the EUreccA 2025 RWE workstream, we set out to consider what might be done to improve RWE and complement or enhance other RWE initiatives – particularly in the context of health technology assessment (HTA), to guide reimbursement and pricing decisions by payers, both public and private. This is discussed further below.

1.2 EUreccA and the RWE workstream

EUreccA 2025 is a think tank, stakeholder forum of academics, policy and payer experts funded by Novartis; it stands for 'European initiative of new Reimbursement and aCCess Approaches'. This think tank advances constructive engagement on the growing challenges of European healthcare between healthcare system stakeholders. It enables sharing of best practice and co-creation of innovative solutions for a sustainable healthcare ecosystem. EUreccA 2025 aims to build trust through transparent discussions and exchanges with the common goal of improving patient outcomes. One of the EUreccA 2025 projects is on the use of RWD/E to support healthcare decision making. The work described in this paper forms part of the EUreccA 2025 RWE workstream and was conducted by our RWE Steering Group (SG) and RWE core project team (see Appendix A for group members).

The project began with scoping work, which generated four key topics salient to identify and address the barriers to optimal RWE use in HTA. Using pragmatic literature searches, stakeholder engagement and case studies, we have identified practical ways in which problems

may be addressed. A central recommendation is that a repository of past decisions using RWE in HTA should be created, which can be accessed speedily using a helpful categorization to facilitate discussions between manufacturers and payers or their agencies. This paper describes the detailed structure of this proposed repository. It is organized around a taxonomy of pairings of questions asked of RWE in HTA and the data sources used to answer the questions. This paper forms part of a group of papers, published as companion pieces in this journal.[1,11-13]

1.3 Making the case for an RWE repository of experience (a taxonomy)

We received a strong steer from the EUreccA 2025 membership that the strengths and weaknesses of RWE should not be itemized without specifying the question being addressed and the source of evidence. This view was validated in our literature reviews and country studies. Paper 1[1] sets out this starting premise of the work. In it, we discuss the view that a repository of experience with data/question pairings would be very helpful to both the companies making submissions to agencies or payers to gain reimbursement, and to the analysts receiving the submissions.[1] This view was based on our literature reviews and validated through our stakeholder engagement.[11] Similar views have been voiced by other bodies since we embarked on the work (e.g. repository of data sources, registry for RWE studies or protocols[15,16]).[17-21] As we knew of no practical resource or database with examples of pairings in practice (RWE questions, RWD sources that can feasibly support these questions and the appropriate methods to use), we set out to create a suitable structure for such a resource – partly to demonstrate how it would work and partly to offer some immediate value. Our belief that a repository is needed is clearly shared by some others, although the form we are suggesting sets it apart somewhat. Some initiatives or associations have now recommended portals for registration of RWE studies[17,20,21], and some have created these portals (ISPOR; the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance [ENCePP]).[15,16] Others have recommended repositories of RWD sources[18] and generation of RWD that is fit for purpose.[18,22,23]

1.4 What format is needed?

The repository should be in a helpful format or structure to support easy and quick access to materials. Importantly, it should be updated and maintained and be shared transparently. We began with a 'taxonomy', which, by definition (Appendix B), is the practice or science of classifying things or concepts via a typology. The classification reflects the fact that there are

many different questions that RWD can be used to answer and also begins to categorize the varying RWD sources. Through our data/question pairings, we link specific examples in practice from published sources to the statistical methods used to address these pairings, and describe the issues that have arisen when the question of concern has been addressed using a particular data source in HTA decision making. Often the latter has not proved possible because no HTA publication has described the issues and how they are factored into a decision. There is a lack of transparency on the part of most HTA agencies on how they have used the analyses provided to arrive at their decisions.

2. Methods used to develop the taxonomy

Our accompanying methods paper provides details of how the taxonomy was generated; we provide a summary here.[11]

To develop the taxonomy, we used targeted literature reviews and stakeholder engagement, initially to generate a preliminary list of pairings in practice compiled according to the typology set out in Paper 1 (see Table 1[1]). We used further searches and stakeholder engagement (including an advisory board) to populate this resource using an iterative process.[11] In addition, staff of the School of Health and Related Research (ScHARR) at the University of Sheffield provided short summaries/commentaries on the issues identified in the literature for a selection of pairings. The pairings we have identified are unlikely to constitute an exhaustive list of all those that have been used (although we have striven for that), and we have certainly not had the resources to populate all of them in the work we have done. We think we have populated sufficient of the pairings to demonstrate the value of a repository of examples structured using the taxonomy as an information source, if it is maintained and curated. Also, we provide examples that HTA practitioners can use to investigate some pairings of interest.

3. Results

3.1   Pairings identified to date

We have identified 270 pairings, i.e. 27 questions addressed using 10 different types of data source (although many more **individual** data sources than 10 are reflected in the literature, and each of these will have their own strengths and limitations).[11] Table 1 sets them out using a matrix of the questions and sources. Using the matrix, readers who are interested in a particular

question can go to it in the table. By following the hyperlinks, they will be brought to specific examples of this question being addressed (in a paper, an HTA submission, or both) and the issues that have arisen. These examples are summarized in Appendix C. For our pairings with commentaries, the user can start with a short summary, covering the study title, study type, the source name (e.g. US renal database), the statistical method used and some limitations or benefits. They can then delve deeper into the associated papers via the links provided and into the methods papers referenced; each specific pairing is supported by high-level RWD/E methods papers or guidelines covering the general issues that relate to the broad type of question and the type of data source (i.e. not focused on a specific pairing) and links to relevant ongoing RWE initiatives retrieved as part of our search.

Where there were gaps in our matrix, we ran tailored searches (TLR Search 2b)[11]; these are denoted as 'S2b' in Table 1. Some gaps still remain (white cells), denoting that we found no examples. We suggest reasons for this in our accompanying methods paper[11] and provide hyperlinks to general guidance (see the header row of Table 1) that can support this pairing until specific examples can be included (Appendix D). These general papers provide details on methodological issues that arise in conducting the targeted analysis using any or all of the ten types of data source. In addition, hyperlinks are provided linking to source-specific guidance at the top of each column (e.g. guidance to using electronic health records [EHRs] or registries summarized in Appendix D). As mentioned, ScHARR provided short commentaries on the issues identified in the literature for 100 of the 270 pairings. These are summarized in Appendix C6 and discuss the strengths and limitations of methods used for specific pairings. Currently, these commentaries are heavily weighted towards effectiveness, reflecting its dominance as a topic for discussion in the literature. A further limitation is that most of the HTA examples included in our taxonomy are from the National Institute for Health and Care Excellence (NICE), which reflects the relative openness of NICE and how that Institute transparently reports explanations of their decisions. The inability to easily understand how other agencies use RWE for particular pairings is itself a telling finding.

3.2 *Practical examples of how to use the taxonomy*

The following text aims to illustrate how the taxonomy can be used in its current form and the information that can be retrieved.

Example 1: A pairing in practice, not necessarily related to an HTA submission

The user has a pairing in mind and wants to find more information about this pairing in practice.

- What is your question? Answer: What is the **effectiveness of a technology?**

- What source are you considering using? Answer: **EHR**

The user would take the following steps:

- **Step 1:** In the **pairing matrix** (Table 1) select the pairing of interest. This takes the user to examples and of this pairing in practice (Appendix C)

- **Step 2:** Note: The intention is to provide some detail so the user can decide if they want to know more about a particular pairing or topic for that pairing, such as:

  - A list of **papers** or **HTA submissions** (if available) for that pairing via links. Note: Table 2 includes three examples (papers) to illustrate the detail provided in Appendix C for the pairing **effectiveness of a technology + EHR**

  - If the user is interested in one paper or the **statistical method** used in that paper, they can find out more via the commentary link (if available). Note: Table 3 includes one commentary example to illustrate the detail provided in Appendix C6 for the pairing **effectiveness of a technology + EHR**

  - The **RWD source methods** & **guidelines** via links included:

    - Selecting **'EHR'** in Table 1 brings the user to Appendix D, Table D3, which includes EHR methods guides such as '*MIT Critical Data's secondary analysis of EHRs* '[24], covering a range of topics including handling missing data, bias and confounding. Note: Table 4 includes three examples (papers) to illustrate the detail provided in Appendix C for the RWD source: **registry**

  - **General RWD/E methods** & **guidelines** via links:

    - Selecting **'General'** in Table 1 brings the user to Appendix D, Table D1. This includes general methods guides such as NICE's real-world evidence framework (launched 23 June 2022)[25], ISPOR's RWE transparency

initiative, Orsini et. al (2020)[26] and the Duke-Margolis Center for Health Policy's RWE white paper (2019).[27] Note: Table 5 includes three examples (papers) to illustrate the detail provided in Appendix D

- **Step 3:** User can decide if this **pairing** is relevant and warrants further investigation:

    - If **YES** (relevant) they can investigate more papers or HTA submissions for this pairing via the **pairing matrix**

    - If **NO** (not relevant) they can go back and select a different pairing from the **pairing matrix**, e.g. effectiveness of a technology + registry

Examples 2 and 3: A pairing in practice, related to an HTA submission

The user has a pairing in mind and wants to find out more information on this pairing in practice but specifically on how it was received by an HTA agency. Firstly, the user should consider their question and source pairing:

- What is your question?

- Example 2: What is the **effectiveness of current standard of care** for the treatment of melanoma?

- Example 3: What is the **effectiveness of a comparator**?

    - What source are you considering using?

- Example 2: SEER **registry**

- Example 3: **Observational** (to facilitate a synthetic control arm)

    Go to Table 1 and in the appropriate matrix cell select the link to examples of this pairing in practice (Appendix C). Here you can check whether there is an appraisal available that discusses this pairing, or alternatively you can see the list of NICE technology appraisals (TAs) included so far in Appendix E. Tables 6 and 7 summarize the type of information that can be retrieved, using Appendix C as an example.

    An example of a NICE TA that used RWE for questions related to using surrogate outcomes to predict long-term clinical effectiveness is TA251 (2012). It assessed dasatinib, nilotinib and standard-dose imatinib as first-line treatment of chronic phase chronic myelogenous leukaemia (CML). Licensing of these treatments was based on randomized controlled trials assessing

complete cytogenetic response and major molecular response at 12 months as primary endpoints. The Committee accepted that people with either a complete cytogenetic response or a major molecular response after 12 months experienced better long-term survival. Oriana et al. (2013) note in their paper the importance of validated surrogate outcomes for policy makers and that this TA highlighted that observational-level evidence may suffice for the new drug to be included in public formularies.[28] The history and reliability of the surrogate seem to have been important factors in its acceptability. Having access to this type of information would be very useful for the user.

Example 4: The user does not have a specific pairing in practice in mind; they would like to understand more about RWE and its use

The user does not have a pairing in mind but wants to find more information about, for example, the varying RWE frameworks, general RWE guidance, a RWE glossary, how RWE was received by HTA agencies regardless of the pairing, RWE checklists, and risk of bias tools. The user would take the following steps:

- In the **pairing matrix** (Table 1) in the section **'other questions (no pairing)'** select **the question you are interested in** and follow the link to the relevant appendices

- Selecting **'General'** in Table 1 brings the user to Appendix D, Table D1, which includes general methods guides such as NICE's real-world evidence framework (as described above)[25]

3.3   The taxonomy can support navigation of the vast evidence base of ongoing initiatives

There is a vast amount of information published on RWE, including useful initiatives, publications, methods, etc. However, this volume of information can stifle the HTA practitioner and decision maker – how can one navigate this in a timely and useful way? The taxonomy includes a set of appendices (summarized in Table 8) that are intended to support navigation.

4. Discussion

4.1   How the taxonomy can be used in future RWE efforts

As noted in Paper 1[1], one of our aims was to create a resource that could be curated, modified if necessary, and added to, as more examples of pairings were identified and as more experience of

the use of already identified pairings was gained. This would be similar to the maintenance of a 'wiki', which could be added to by the research community (collaboratively edited). For example, a 'submitter' would write up their pairing example and a curator would review to ensure that it is of acceptable standard before adding to the repository. It is not intended to replace the work that is going on by others to develop methods and good practice (see Appendix D for examples). Instead, it is intended to complement these methods and facilitate their use, as well as provide some basis for agreeing when evidence is likely to be 'good enough' for a key decision despite (inevitably) falling short of perfection, and to allow both agency and manufacturer to understand before a submission the likely consequences of a particular approach.

4.2   Limitations

The intention of the summaries in the appendices is to provide some detail on how a pairing was received. The user can then use this as a basis to identify more pairings, guidelines, methods, etc. A limitation of this work is that we have only included some examples of NICE TAs covering some pairings; examples of NICE TAs are not included for all pairings. We did not have the time nor the resources to incorporate everything on our first attempt, and the aim would be to add more examples as the taxonomy is curated in the future. It should be noted that Ciani et al. (2021) reported on the validity of surrogate endpoints (question) and their impact on coverage decisions in several countries including England (NICE).[29] For this question alone, they screened 291 HTA reports from NICE, which gives an indication of the extent of work required to compile this information.

We think it is certain that the classification can be improved and extended, and we are aware that for many of the categories we have defined there are few sources of useful guidance. The scale of the literature on the proper use of RWE and its rate of growth means that a single source set of interlinked documents, initially incomplete, such as we have created will quickly go out of date. To be useful, the taxonomy must become a living document, fed material largely by the community that benefits from it.

What we have offered here, with some trepidation, is a suggestion for a core classification. The literature base is so large that we have not been able to review it all and, in particular, our restriction mainly to English language sources may mean we have missed something important. In addition, we provide one approach to organizing information, but there could be other ways.

We have attempted to guard against these issues by consulting widely both inside and outside the EUreccA 2025 initiative with people of many perspectives, nationalities and languages. Limitations remain, but we hope that the work we have done will provide a firm base from which a bigger, more valuable, living resource can be developed, one that could enable better use of RWD and RWE to get effective treatments more quickly to the patients who can benefit.

4.3   *Taxonomy summary and potential impact*

A major stumbling block to the fuller acceptance and use of RWE in HTA and reimbursement decision making for pharmaceuticals lies in a lack of trust on the part of decision makers that analyses based on RWD can be relied upon and not be manipulated to give results biased in favour of a submitting company. This is reinforced by the difficulties in anticipating the nature and importance of particular limitations in a given data source for the particular purpose for which it is intended. This is certainly the case in using RWE to estimate the relative effectiveness of interventions, despite a very large and rapidly growing literature guiding practice.

Recently, there have been calls for much more openness in the use of RWE[17,18,30,31] and for an open resource that allows examination of the way RWE is used. This mirrors the current system of registration and publication of clinical trials protocols and results. However, this is difficult to achieve when using RWE in HTA. Not only does it require an organization to be funded to take on maintaining and curating the resource – a very significant task – but it also requires a structure and content for the resource that will ensure it is useful. Guided by a wide group of colleagues and what we have found in the literature, we have developed a tool based on the **particular** questions that RWE is used to answer, paired with the **particular** (type and detailed) source(s) of RWD that may be and (where we found examples) have been used. We hope this will enable users to do several things:

- First, to identify the issues with which they will have to grapple in using their data sources to answer their question

- Second, to easily identify relevant, high-quality methodological guidance

- Third, to learn what has been the experience of others faced with similar challenges in practice – i.e. how agencies have responded in the past to receiving submissions tackling similar questions

Our work is not complete, but we believe it is sufficient to demonstrate the value of such a taxonomy and information source if it is completed and curated as a 'wiki' by the community that would use it. As we mention, to be useful, the taxonomy has to become a living document, fed material largely by the community that benefits from it. Future work involves finding an RWE organization that is best placed to host, champion, curate, maintain and improve the taxonomy. The authors are in ongoing discussions to support this. In addition, digital technology such as artificial intelligence and natural language processing could potentially help identify and categorize RWE publications/reports into our pairing categories.

Note: Not all guidelines, technology appraisals, etc. included in the appendices are included here as references. Instead, a hyperlink to that document has been included in the appendix.

References

1. Akehurst R, Murphy LA, Solà-Morales O, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Paper 1: Using real-world data in the health technology assessment of pharmaceuticals: strengths, difficulties and a pragmatic way forward. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2022.

2. Adapt Smart. Accelerated Development of Appropriate Patient Therapies. 2020; http://adaptsmart.eu/webinar-implementing-real-world-data-to-make-adaptive-pathways-work/. Accessed 4 August 2022.

3. CanREValue. Developing a framework for incorporating real-world evidence into drug funding decisions. 2020; https://cc-arcc.ca/wp-content/uploads/2020/03/CanREValue-Methods-WG-Interim-Report-2020.pdf. Accessed 4 August 2022.

4. EFPIA. EFPIA's Position on the Use & Acceptance of Real World Evidence by International Markets. 2021; https://www.efpia.eu/media/602956/efpia-rwe-position-paper_aug2021.pdf. Accessed 4 August 2022.

5. GetReal Institute. GetReal: Facilitate the adoption and implementation of RWE in health care decision-making in Europe. 2021; https://www.getreal-institute.org/. Accessed 4 August 2022.

6. ISPE. ISPE'S POSITION ON REAL-WORLD EVIDENCE (RWE). 2020; https://pharmacoepi.org/pub/?id=136DECF1-C559-BA4F-92C4-CF6E3ED16BB6. Accessed 4 August 2022.

7. ISPOR. ISPOR's Real-World Evidence Transparency Initiative: Creating a Road Map That Builds Trust. 2020; https://www.ispor.org/heor-resources/news/2020/10/05/ispor-real-world-evidence-transparency-initiative-creating-a-road-map-that-builds-trust. Accessed 4 August 2022.

8.   ISPOR, Duke-Margolis Center for Health Policy, National Pharmaceutical Council. RWE Transparency initiative 2021; https://www.ispor.org/strategic-initiatives/real-world-evidence/real-world-evidence-registry/?utm_medium=press_release&utm_source=public&utm_campaign=general_ispor&utm_content=press_release_oct26&utm_term=rwe_registry Accessed 4 August 2022.

9.   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. *Journal of comparative effectiveness research.* 2021;10(9):711-731.

10.   RWE4Decisions. RWE4Decisions: Defining a new vision for the use of RWE. 2021; https://rwe4decisions.com/. Accessed 4 August 2022.

11.   Murphy LA, Akehurst R, Solà-Morales O, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Paper 2: The real-world evidence workstream in EUreccA 2025: How the task was addressed *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2022.

12.   de Pouvourville G, Cunningham D, Fricke FU, et al. Paper 4: Across-country variations of Real-World Data and Evidence for drugs. A five-European-country study *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2022.

13.   Solà-Morales O, Akehurst R, Murphy LA, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Paper 5: Data governance for real-world data management: optimizing its use in HTA and regulatory decision-making. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2022.

14.   Murphy LA, Akehurst R, Solà-Morales O, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Paper 3: Structure and content of a taxonomy to support the use of real-world evidence by health technology assessment practitioners and healthcare decision-makers. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2022.

15. ISPOR, Duke-Margolis Center for Health Policy, (NPC) NPC. Real-World Evidence Registry. 2022; https://osf.io/registries/rwe/discover. Accessed 4 August 2022.

16. EnCeePP (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance). Real-World Evidence Study Registry. 2022; https://www.encepp.eu/encepp/viewResource.htm?id=36064. Accessed 4 August 2022, 2022.

17. 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. *International journal of technology assessment in health care.* 2020:1-10.

18. 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. *International journal of technology assessment in health care.* 2019;35(4):346-350.

19. Beaulieu-Jones BK, Finlayson SG, Yuan W, et al. Examining the Use of Real-World Evidence in the Regulatory Process. *Clinical pharmacology and therapeutics.* 2020;107(4):843-852.

20. Hampson G, Towse A, Dreitlein WB, Henshall C, Pearson SD. Real-world evidence for coverage decisions: opportunities and challenges. *Journal of comparative effectiveness research.* 2018;7(12):1133-1143.

21. Pearson SD, Dreitlein WB, Towse A, Hampson G, Henshall C. A framework to guide the optimal development and use of real-world evidence for drug coverage and formulary decisions. *Journal of comparative effectiveness research.* 2018;7(12):1145-1152.

22. O'Leary CP, Cavender MA. Emerging opportunities to harness real world data: An introduction to data sources, concepts, and applications. *Diabetes, obesity & metabolism.* 2020;22 Suppl 3:3-12.

23. Schneeweiss S, Glynn RJ. Real-World Data Analytics Fit for Regulatory Decision-Making. *American journal of law & medicine.* 2018;44(2-3):197-217.

24. MIT Critical Data. *Secondary analysis of electronic health records* Springer International Publishing; 2016.

25. National Institute for Health and Care Excellence (NICE). NICE real-world evidence framework. 2022; https://www.nice.org.uk/corporate/ecd9/chapter/overview. Accessed 11 October 2022.

26. Orsini LS, Monz B, Mullins CD, 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. *Pharmacoepidemiology and drug safety.* 2020;29(11):1504-1513.

27. Mahendraratnam N, Eckert L, Mercon K, et al. White Paper. Understanding the Need for Non-Interventional Studies Using Secondary Data to Generate Real-World Evidence for Regulatory Decision Making, and Demonstrating Their Credibility. 2019; https://healthpolicy.duke.edu/publications/understanding-need-non-interventional-studies-using-secondary-data-generate-real-world. Accessed 07 October 2022.

28. Oriana C, Martin H, Toby P, et al. Complete cytogenetic response and major molecular response as surrogate outcomes for overall survival in first-line treatment of chronic myelogenous leukemia: a case study for technology appraisal on the basis of surrogate outcomes evidence. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2013;16(6):1081-1090.

29. Ciani O, Grigore B, Blommestein H, et al. Validity of Surrogate Endpoints and Their Impact on Coverage Recommendations: A Retrospective Analysis across International Health Technology Assessment Agencies. *Medical decision making : an international journal of the Society for Medical Decision Making.* 2021;41(4):439-452.

30. 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; https://www.ispor.org/docs/default-source/strategic-initiatives/improving-transparency-in-non-interventional-research-for-hypothesis-testing_final.pdf?sfvrsn=77fb4e97_6.

31. 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. *International journal of technology assessment in health care.* 2019;35(4):346-350.

Tables

Table 1: Question and source 'pairing matrix'

QUESTION	SOURCE									
	General (Appendix D1)									
(See relevant Appendices for examples of pairings in practice)	EHR (Appendix D3)	Registry (Appendix D2)	Observational study (Appendix D4)	Post-trial study	PCTs (Appendix D5)	Survey	Social media, consumer device (Appendix D7)	EAP	Claims/administrative data (Appendix D6)	Biobanks/genomic data (Appendix D8)
Effectiveness of a technology										
What is the effectiveness of a technology, does this change over time?	Appendix C1	Appendix C1	Appendix C1	Appendix C1	Appendix C1	**	Appendix C5 (S2b)		Appendix C1	Appendix C1
What is the effectiveness in a subgroup, does this change over time?	Appendix C1	Appendix C1	Appendix C1					*	Appendix C1	Appendix C5 (S2b)
What is the effectiveness of its comparator(s), does this change over time?	Appendix C1	Appendix C1	Appendix C1	Appendix C1	Appendix C1		Appendix C1		Appendix C1	Appendix C5 (S2b)
What is the relative treatment effect, does this change over time?	-	Appendix C1	Appendix C1		Appendix C5 (S2b)		Appendix C5 (S2b)		Appendix C1	Appendix C5 (S2b)
What are the main modifiers of effect (e.g. covariate which are predictive of effect)?	Appendix C1	Appendix C1	Appendix C1		Appendix C5 (S2b)		Appendix C5 (S2b)		Appendix C1	Appendix C5 (S2b)

QUESTION	SOURCE									
	General (Appendix D1)									
(See relevant Appendices for examples of pairings in practice)	EHR (Appendix D3)	Registry (Appendix D2)	Observational study (Appendix D4)	Post-trial study	PCTs (Appendix D5)	Survey	Social media, consumer device (Appendix D7)	EAP	Claims/administrative data (Appendix D6)	Biobanks/genomic data (Appendix D8)
Are there biomarkers which are surrogates for mortality or how patients feel?	† See General guidance (Appendix C1): Duke (2020) includes details on developing surrogate endpoints, AHRQ (Book, 2019) includes details of surrogate markers (registries, EHRs). Ciani et al. (2021) reported on validity of surrogate endpoints (question) and their impact on coverage decisions in several countries including England (screened 291 NICE HTA reports) [15].								Appendix C1	† Appendix C5 (S2b)
Patient and the patient experience										
What is the patient's burden of illness?	Appendix C3	Appendix C5 (S2b)	Appendix C5 (S2b)	**	Appendix C5 (S2b)	Appendix C3	Appendix C5 (S2b)	*	Appendix C3	Appendix C3
What are the key domains for ADL, QoL and other PROMs and PREMs?										
How, and when, is the disease diagnosed?	Appendix C3	Appendix C5 (S2b)	Appendix C5 (S2b)			Appendix C5 (S2b)	Appendix C3		Appendix C3	Appendix C3
What are the characteristics of patients with this condition?	Appendix C3	Appendix C3	Appendix C5 (S2b)			Appendix C5 (S2b)	Appendix C3		Appendix C5 (S2b)	Appendix C3
What are causes, predictors or risk factors for the disease? Identify patients, estimate sample size	Appendix C3	Appendix C3	Appendix C3			Appendix C3	Appendix C3		Appendix C3	Appendix C3
TREATMENT: When are patients typically treated? When does treatment	Appendix C3	Appendix C5 (S2b)	Appendix C5 (S2b)		Appendix C5 (S2b)	Appendix C5 (S2b)	Appendix C5 (S2b)		Appendix C3	Appendix C5 (S2b)

QUESTION	SOURCE									
	General (Appendix D1)									
(See relevant Appendices for examples of pairings in practice)	EHR (Appendix D3)	Registry (Appendix D2)	Observational study (Appendix D4)	Post-trial study	PCTs (Appendix D5)	Survey	Social media, consumer device (Appendix D7)	EAP	Claims/administrative data (Appendix D6)	Biobanks/genomic data (Appendix D8)
start (start criteria)?										
What is the pathway of care?										
What current treatments are available?										
What criteria are used to stop treatment?										
How is a product used in practice e.g. support clinical decisions, guideline adherence	Appendix C3	Appendix C3	Appendix C3		Appendix C3	Appendix C3				
How is a product used in practice e.g. Dose, frequency, adherence, persistence	Appendix C3		Appendix C5		Appendix C3	Appendix C5	Appendix C5		Appendix C3	Appendix C3
What is the impact of a particular technology on ADL and QoL?	Appendix C5 (S2b)	Appendix C5 (S2b)	Appendix C5 (S2b)		Appendix C3	Appendix C3	Appendix C5 (S2b)		Appendix C5 (S2b)	Appendix C5 (S2b)
Adverse events										
What are the AEs/side effects experienced by patients receiving a treatment?	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2	Appendix C2

QUESTION	SOURCE									
	General (Appendix D1)									
(See relevant Appendices for examples of pairings in practice)	EHR (Appendix D3)	Registry (Appendix D2)	Observational study (Appendix D4)	Post-trial study	PCTs (Appendix D5)	Survey	Social media, consumer device (Appendix D7)	EAP	Claims/administrative data (Appendix D6)	Biobanks/genomic data (Appendix D8)
What are the additional safety concerns compared with comparators?										
What is the time profile of AEs?		Appendix C2								
What are the causes, predictors, risk factors for AEs, time to event, covariates	Appendix C2	Appendix C5 (S2b)	Appendix C3				Appendix C2	Appendix C2	Appendix C2	Appendix C2
Costs, resource use										
What is the use of health services resources over time?	Appendix C4	Appendix C4	Appendix C4		Appendix C5 (S2b)	Appendix C5 (S2b)			Appendix C4	
What are the covariates with resource use?										
What are costs associated with treatments and resource use?	Appendix C4	Appendix C4	Appendix C4		Appendix C4	Appendix C4		*	Appendix C4	
What is the budget impact?										
Is the technology used according to indication and potential										

QUESTION	SOURCE									
	General (Appendix D1)									
(See relevant Appendices for examples of pairings in practice)	EHR (Appendix D3)	Registry (Appendix D2)	Observationa l study (Appendix D4)	Post-trial study	PCTs (Appendix D5)	Survey	Social media, consumer device (Appendix D7)	EAP	Claims/administrativ e data (Appendix D6)	Biobanks/genomi c data (Appendix D8)
reimburseme nt limitations?										
Other questions (no pairing)										
What RWE frameworks and initiatives currently exist? (Appendix F3)										
What RWE checklists and risk of bias tools are available? (Appendix F1 and F2)										
Are there examples of NICE Technology Appraisals where RWE has been used previously? (Appendix E)										
Do we have a glossary of terms? (Appendix B)										

Key: ADL, activities of daily living; AE, adverse events; BOI, burden of illness; EAP, Expanded Access Programme; EHR, electronic health records; Obs, observational; PCT, pragmatic controlled trial; PROMs, patient-reported outcome measures; PREMs, patient-reported experience measures; QoL, quality of life; RU, resource use.

Notes: Where there are gaps (white cells) of specific pairings, we suggest high-level RWE guidance documents (links are included). IMI-Get Real covers PCTs, see PCTs and General guidance. *, Results from EAP are not typically published. **, This is not typically a source for this question.

Table 2: Examples of pairings in practice

What is the effectiveness of a technology, how does this change over time? (results from Search 2)					
Use (question)	Source	Title/study type/statistical method(s) – statistical question	Source names	Limitations/benefits (example)	Link to paper / commentary
Effectiveness	EHR	**Comparing Effectiveness of Dynamic Treatment Strategies Using EHRs: An Application of Parametric g-Formula to Anemia Management Strategies**/review of methods/parametric G-formula – control for unmeasured confounding in observational studies	US Renal Data System	The G formula should be carried out with IP weighing to assess the g-null paradox or, when possible, doubly robust estimates should be used to combine the benefits of both methods	https://doi.org/10.1111/1475-6773.12718 (Zhang, 2018) Commentaries: Appendix C6
Effectiveness – does the intervention work	EHR	**Analytical Methods for a Learning Health System: 3. Analysis of obs Studies**/methods guide/regression, approaches ITS analyses, IVs, PSMs – Address bias	Mentions Medicaid and Medicare	The paper describes how analytical methods for individual-level EHD, including regression approaches, ITS analyses, IVs and PSMs, can be used to address the question of whether the intervention 'works'. Limitations of each also discussed	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5982993/pdf/egems-5-1-252.pdf (Stoto, 2017) Commentaries: Appendix C6
Effectiveness – long-term effectiveness, covariates of effectiveness (socioeconomic)	EHR	**Effectiveness of varenicline versus NRT on long-term smoking cessation in primary care: a prospective cohort study of EMRs**/application of a method/multivariable logistic regression, PSM and IV analyses – Confounding	CPRD	A particular strength of this study was the use of three different analytical methods to estimate the effectiveness of varenicline. The propensity score balanced the treatment groups' observed baseline characteristics, and produced similar findings to the multivariable adjusted regression. IV analyses used naturally occurring variation in GPs' prescribing which, if its assumptions hold, is robust to unmeasured residual confounding of exposure–outcome relationship, including confounding by indication	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5837420/ (Taylor, 2017) Commentaries: Appendix C6

Note: Full set of examples are included in Appendix B.

Table 3: One example of a ScHARR commentary for Zhang (2018)

Title	Comparing the Effectiveness of Dynamic Treatment Strategies Using Electronic Health Records: An Application of the Parametric g-Formula to Anemia Management Strategies
Authors	Y. Y. Zhang, J. G.: Thamer, M.: Hernán, M. A.
Country	US
Year	2018
Publication Type	☒ Journal article ☐ Report ☐ Other (please specify:_____)
Drug / Technology	Dialysis
Disease Area	ESRD – CHF, IHD
Study type	☐ Review of one or several methods ☐ Method guide ☐ Methods paper (discussion of a method) ☒ Case study application of a method ☐ Other (please specify:_____)
Study aim	To compare the effectiveness of dynamic anaemia management strategies by applying the parametric g-formula to electronic health records.

Source(s)*	☐ EHR ☐ Registry ☒ Observational ☐ Administrative data ☐ Claims data ☐ Pharmacy data	☐ Post-trial study, early trial study ☐ Survey ☐ Social media ☐ Consumer device ☐ Biobank ☐ Other (please specify)
Source(s) names / case studies	US Renal Data System	
Question(s)*	☒ Effectiveness – and safety ☐ Effectiveness, comparative effectiveness ☐ Effectiveness, long-term effectiveness ☐ Effectiveness, in a specific subgroup ☐ Adverse events / safety - death ☐ Diagnosis - Patterns of symptoms before diagnosis	☐ Patient characteristics ☐ Product use in practice, e.g. support clinical decisions, throughput volumes, adherence to guidelines, discharge status ☐ Cost, resource use ☐ Bench-marking against RCTs ☐ Other (please specify)

Methods / discussed and reason used	G formula (Young, Hernan, and Robins 2014).
Summary of benefits and limitations of statistical methods	• **What are the main recommendations?** The G formula should be carried out with inverse probability (IP) weighing to assess the g-null paradox, or, when possible, doubly robust estimates should be used to combine the benefits of both methods. • **Is HTA decision making discussed?** No
Checklist	• Rating from screening/title abstract = 6 • Main reason for this score = 15 methods based on LIT SEARCH - unmeasured confounders in comparative observational studies • Rating from full review (i.e. do you think the above rating should change?) = 6 • Main reason for this change in score = no change
Link to publication	https://doi.org/10.1111/1475-6773.12718

Note: n = 100 included in Appendix C6. *Refer to protocol for full list of sources and question.

Table 4: Examples of specific 'source' methods guides – Methods guides for registries

Initiative/agency	Title	Details	Link
FDA (2021)	*Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry (Draft Guidance)*	*Provides guidance on using registry data to support regulatory decisions, relevance of registry data, reliability of registry data, considerations when linking a registry to another registry or another data system and considerations for regulatory review.*	[Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry	FDA](#)
AHRQ (2013, 2018, 2019)	*Book: Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition (2013); Addendum I - 21st Century Patient Registries (patient centric registries) (2018); Addendum II - Tools and Technologies for Registry Interoperability (2019)*	*Includes Chapter 5: PROs.* *Chapter 13: Analysis, Interpretation, and Reporting of Registry Data To Evaluate Outcomes (provides nice examples – example of a PRO use in a registry, surrogate markers)* *Chapter 6, 15-18: Discusses how to link registries with EHRs or other sources. Examples included.* *Addendum II includes strengths, weaknesses of EHRs, claims data, genomic data, common data models. Includes chapter on data sources.* *Comprehensive resource, also includes sections on creating registries, legal and ethical considerations, operating registries (data collection and quality assurance)*	https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition/research
AHRQ (2018) *Registries*	*White Paper: Managing Missing Data in Patient Registries*	*Includes details on reasons for missing data, how to minimize missing data, methods to account for missing data (complete case analysis strategy, single and multiple imputation, inverse probability weighting, maximum likelihood methods), reporting guidelines and sensitivity analysis. Useful for definitions in Taxonomy.*	https://effectivehealthcare.ahrq.gov/products/registries-guide-4th-edition/white-paper-2016-2

Note: Full list of examples included in Appendix C.

Table 5: Examples of 'general' RWD/E methods guides – General RWE guidance documents

Initiative/agency	Title	Details	Link

NICE (2022)	*Real-world evidence framework (Launched 23 June 2022)*	• *NICE Framework which clearly describe best-practices for the planning, conduct, and reporting of real-world evidence studies*
		• *improve the transparency and quality of real-world evidence used to inform NICE guidance*
		• *improve committee trust in real-world evidence studies*
		• *ensure real-world evidence is used where it helps to:*
		    ○ *reduce uncertainties*
		    ○ *improve recommendations*
		    ○ *speed up access of patients to new effective interventions.*
			https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

NICE (2021)	*NICE strategy 2021 to 2026*	*Discusses 'the need to integrate real-world data into our evaluation processes to inform rapid but robust decisions'.*	*https://static.nice.org.uk/NICE%20strategy%202021%20to%202026%20-%20Dynamic,%20Collaborative,%20Excellent.pdf*
NICE-2020	*NICE methods of health technology evaluation: the case for change*	**Announcement:** *NICE revisiting methods, process includes refreshing and clarifying:* • *Emphasis on role of comprehensive evidence base, including non-RCTs & RWE, circumstances in which different types of evidence have strengths or limitations* • *Additional guidance on use of RCT and non-RCT evidence, assessment and reporting of study quality, risk of bias and confounding, and presenting evidence.*	https://www.nice.org.uk/news/article/nice-s-methods-of-technology-evaluation-presenting-a-case-for-change
RWE Transparency initiative (2021)	*RWE registry to register study protocols etc.*	• *RWE Transparency Initiative is a partnership between ISPOR, the Duke-Margolis Center for Health Policy, and the National Pharmaceutical Council.* • *New RWE Registry Developed From the RWE Transparency Initiative launched* • *The registry, hosted on Open Science Framework (OSF), developed and maintained by Center for Open Science, provides researchers with a platform to register their study protocols before they begin work. Using open, centralized workflows enhances collaboration and facilitates the transparency needed to elevate the trust in the study results.*	https://www.ispor.org/strategic-initiatives/real-world-evidence/real-world-evidence-registry/?utm_medium=press_release&utm_source=public&utm_campaign=general_ispor&utm_content=press_release_oct26&utm_term=rwe_registry

Note: Full list of examples included in Appendix C.

Table 6: Pairing example that relates to a NICE Technology Appraisal (1)

Question	What is the effectiveness of current standard of care for the treatment of melanoma?
Source	SEER registry
TA title	*Vemurafenib Melanoma*
TA number (year)	TA269 (2015)
Feedback	Committee accepted idea of using external evidence to support decision making, however, each of preferred sources disputed by ERG and Committee. Data from an observational study by Balch et al. preferred over SEER registry data as Balch data allowed for adjustment according to staging of disease, and utility values from another study were preferred over standard gamble study.
Was RWE accepted?	Committee ultimately accepted RWE **but not the original sources**. Important to check for alternative RWE sources.
Link to TA	https://www.nice.org.uk/guidance/TA269
Link to commentary / further details	Not available, however in this case please see Bullement et al., (2020) for further details of the TA: https://pubmed.ncbi.nlm.nih.gov/32646531/. Bullement (2018) A review and validation of overall survival (OS) extrapolation in HTAs of cancer immunotherapy by the NICE: how did the initial best estimate compare to trial data subsequently made available? https://www.tandfonline.com/doi/full/10.1080/13696998.2018.1547303
Link to general methods, examples	E.g. CanREValue Collaboration Methods Working Group Progress Report on Real World Survival Data https://cc-arcc.ca/wp-content/uploads/2020/03/CanREValue-Methods-WG-Interim-Report-2020.pdf E.g. AHRQ Managing Missing Data in Patient Registries https://effectivehealthcare.ahrq.gov/products/registries-guide-4th-edition/white-paper-2016-2

Key: ERG, Evidence Review Group; TA, technology appraisal.

Table 7: Pairing example that relates to a NICE Technology Appraisal (2)

Question	What is the effectiveness of a comparator using a synthetic control arm? (facilitate a comparison with a pivotal single arm trial of avelumab)
Source	Retrospective observational studies (EHR, registry data)
TA title	*Avelumab for treating metastatic Merkel cell carcinoma*
TA number (year)	TA517 (2018)
Feedback	ERG and Committee criticised the methods used to analyse the data from the observational studies, requesting the use of matching methods to adjust for differences in patient characteristics and noting the retrospective design which may lead to date of publication bias.
Was RWE accepted?	Committee ultimately accepted RWE based on the observational data to inform decision making.
Link to TA	https://www.nice.org.uk/guidance/ta517
Link to commentary / further details	Not available, however in this case please see Bullement et al., (2020) for further details of the TA: https://pubmed.ncbi.nlm.nih.gov/32646531/
Link to general methods, examples	E.g. Synthetic and External Controls in Clinical Trials – A Primer for Researchers (Thorlund, 2020) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7218288/pdf/clep-12-457.pdf E.g. FDA (2021) Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry (Draft Guidance) https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-registries-support-regulatory-decision-making-drug-and-biological-products
Link to further papers on this topic	Anderson (2018) A review of NICE appraisals of pharmaceuticals 2000-2016 found variation in establishing comparative clinical effectiveness: https://pubmed.ncbi.nlm.nih.gov/30236484/

Key: ERG, Evidence Review Group; TA, technology appraisal.

Table 8: What information is included to support the HTA practitioner and decision maker and why

Appendix	Details	Why is this useful?
A	A glossary of taxonomy terms and methods terms	To facilitate establishing a common terminology
B	Pairings in practice • B1 Effectiveness • B2 Safety and adverse events • B3 The patient • B4 Economic burden and costs • B5 Further pairings to fill gaps • B6 ScHARR commentaries (n=100 pairings in practice with details of methods used)	To facilitate signposting of information – Includes links to summaries and commentaries of examples in practice to get the practitioner or decision maker started
C	Methods guides • C1 General RWE guidance documents • C2 Registry guidance documents • C3 Electronic Health Record guidance documents • C4 Observational data guidance documents • C5 Pragmatic Controlled Trials guidance documents • C6 Claims/administrative data guidance documents • C7 Social media/consumer device guidance documents • C8 Biobank/genomic data guidance documents	To facilitate signposting of information – Includes links to guides to get the practitioner or decision maker started
D	NICE Technology Appraisal RWE examples	An example of how we can combine information from previous HTAs
E	44 RWE checklists and x risk of bias (RoB) assessment tools	To facilitate signposting of information – Links to checklists and RoB tools to facilitate the practitioner and decision maker
F	DOAC case study DOAC pairings in practice	To facilitate signposting of information – Includes links to summaries and commentaries of examples in practice to get the practitioner or decision maker started

Key: DOAC, direct oral anticoagulation; HTA, health technology assessment; RoB, risk of bias; RWE, real-world evidence.