BMJ Open

Understanding the use of real-world data in appraisals of cancer therapy: a protocol for data extraction

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055985
Article Type:	Protocol
Date Submitted by the Author:	02-Aug-2021
Complete List of Authors:	Kang, Jiyeon; London School of Hygiene and Tropical Medicine Faculty of Public Health and Policy, Department of Health Services Research and Policy; Cairns, John; London School of Hygiene and Tropical Medicine, Department of Health Services Research and Policy
Keywords:	ONCOLOGY, Health economics < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

Understanding the use of real-world data in appraisals of cancer therapy: a protocol for data extraction

Jiyeon Kang[1], John Cairns[1]

[1]Department of Health Service Research and Policy,

Faculty of Public Health and Policy,

London School of Hygiene and Tropical Medicine, London, UK

Corresponding author: Jiyeon Kang (Jiyeon.Kang@lshtm.ac.uk)

15-17 Tavistock place, London, WC1H 9SH, UK

Word count: 4,169

Keywords: Real-world data (RWD), Health Technology assessment (HTA), economic evaluation, oncology medicine, National Institute for Health and Care Excellence (NICE)

ABSTRACT

Introduction

Due to the limitations of relying on randomised controlled trials, the potential benefits of real-world data (RWD) in enriching evidence for health technology assessment (HTA) are highlighted. Despite increased interest in RWD, there is limited systematic research investigating how RWD has been used in HTA. The main purpose of this protocol is to extract relevant data from National Institute for Health and Care Excellence (NICE) appraisals in a transparent and reproducible manner in order to determine how NICE has incorporated a broader range of evidence in the appraisal of oncology medicines.

Methods and analysis

The appraisals issued between January 2011 to May 2021 are included following inclusion criteria. The data extraction tool newly developed for this research includes the critical components of economic evaluation. The information is extracted from identified appraisals in accordance with extraction rules. The data extraction tool will be validated by a second researcher independently. The extracted data will be analysed quantitatively to investigate to what extent RWD has been used in appraisals. This is the first protocol to enable data to be extracted comprehensively and systematically in order to review the use of RWD.

Ethics and dissemination

This study is approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine on 14 November 2019 (17315). Results will be published in peer-reviewed journals.

STRENGTHS AND LIMITATIONS OF THIS STUDY

Strengths

- This protocol enables data to be extracted in a transparent and systematic manner for the study of how RWD has been used in NICE appraisals.

- It includes all the different ways an economic evaluation might use RWD.

- This study facilitates systematic understanding of the use of RWD in NICE appraisals over the last 10 years.

Limitations

- Since it is focussed on cancer, the methods and eventually the findings are to some extent cancer-specific.

- The extraction protocol is not fully applicable to the practice of other HTA bodies as much of the protocol reflects the NICE appraisal process. However, the protocol could be modified to reflect the HTA context in different countries.

INTRODUCTION

In the last few years, interest in real-world data (RWD) has grown in health care decision-making (1). Health Technology Assessment (HTA) requires valid and reliable information for the systematic evaluation of health technology. Randomised controlled trials (RCTs) have mainly provided the information (2). However, it is challenging to meet all information needs from RCTs since the new generation of therapies pose several assessment challenges. For example, when treatment options are expanding rapidly, it is increasingly unlikely that there are RCTs featuring all the relevant comparators. Furthermore, the traditional design of RCTs is possibly less appropriate for new technologies such as those targeting rare genetic mutations or where there may be ethical issues with control arms. Another barrier to obtaining all the relevant information for HTA from RCTs concerns the extrapolation of survival. RCTs observe the clinical outcome for a certain period. Extrapolation is required in order to incorporate the survival data from RCTs in the health economic model (3). It is more challenging to identify the most appropriate extrapolation the shorter the duration of the trial. If survival data from RCTs are based on a very limited observation period, the extrapolation of the survival curve is likely to fail to predict the long-term effect (4). Moreover, RCTs tend

to include strictly controlled populations.

The potential benefits of RWD in enriching evidence for HTA are highlighted by the limitations of relying on RCTs (5). This research focuses on the use of RWD in HTA by the National Institute for Health and Care Excellence (NICE). As a leading HTA agency, NICE has achieved an international reputation for rigorous development and application of scientific methods to appraise new health technologies to provide its decisions with robust and fair justification (6). More importantly, NICE is noted for the transparency of its processes, responsiveness to change, and commitment to using the best available evidence (7). The evidence is structurally well-documented enough to find the key information and available on the NICE website. Therefore, review of these appraisals can provide comprehensive information on the evidence used for decision-making. In April 2020, NICE signalled its intention to integrate broader types of data in developing NICE guidance (8). Although it is primarily a statement of intent, it is not a new development in NICE practice since NICE already incorporates a diverse range of published scientific evidence when developing its guidance on health technologies. For example, UK audit data (TA255, 2012), Hospital Episode Statistics (TA559, 2018) and registry data such as the Edinburgh Ovarian Cancer Database (TA598, 2019), Surveillance, Epidemiology, and End Result program (TA562, 2019) have been used in the development of NICE technology appraisal (TA) guidance. While a wide range of data are already used in NICE guidance, there is limited understanding regarding how and where RWD has been used, and in which circumstances RWD is accepted as relevant. Research is required to investigate systematically patterns in the use of RWD and to understand the driving forces behind its use in NICE appraisals.

Several researchers have reviewed practice across HTA bodies (9,10) or reported the use of

RWD in HTA (11). However, little systematic research has been conducted. Important information is missing such as how they included literatures without selection bias, which parts of evidence were reviewed, whether they have clearly defined RWD and justified or explained why this definition is relevant and how different HTA systems were compared given their different practices. Roberts et al. addressed the potential role of RWD in bridging the evidence gaps (12). However, they illustrate the use of RWD with a few examples, rather than providing a fuller picture of current practice when using RWD. Bullement et al. recently reviewed how RWD informed single technology appraisals of cancer drugs in NICE (13). Although this study follows a more systematic approach to review the use of RWD, it does not fully explain how the data were extracted and what criteria were used to judge the use of data. As the process of reviewing appraisals is not clear enough, it is unclear whether the information presented provides a full picture of the use of RWD. Bullement et al. included 113 STAs issued between April 2011 and October 2018. As interest in RWD is increasing over time, it may miss relevant information from recent years. Therefore, the careful development of this extraction protocol is required to help extract the data systematically from appraisals, increase the reliability of the results of the analysis and permit a more detailed description of the use of RWD and analysis of factors influencing its use.

A protocol is required to ensure the consistency of data extraction so that the risk of unsystematic data collection is reduced. The main purpose of this protocol is to extract data from NICE appraisals in a transparent and reproducible manner to answer, “how has NICE incorporated a broad range of evidence in the appraisal of oncology medicines.” Without proper justification and operational rules, the data are likely to be collected erroneously, with a risk of biasing the analysis. The extracted data are expected to be objective and less biased.

With such data, the analysis can provide more robust answers to questions regarding how RWD has been used in NICE technology appraisals. Furthermore, this protocol facilitates the development of a rich dataset which can highlight not just where RWD has been used but also what types of evidence have been used in the HTA process in line with NICE's interest in incorporating a broad range of evidence. The data can be analysed to answer the research questions in depth.

METHODS AND ANALYSIS

Figure 1 simplifies the process of the research. NICE appraisal documents are identified following inclusion criteria. The information is extracted from identified appraisals in accordance with extraction rules. The extraction tool includes general appraisal information and appraisal-specific information such as characteristics of the main clinical evidence and the economic evaluation model.

Definition of RWD

Before extracting information about the use of RWD in NICE technology appraisals, a definition of RWD is clearly required. RWD is an umbrella term which covers broad categories of data. Although RWD is increasingly addressed in the literature, there is no consensus over the definition. In the NICE Evidence Standards Framework for Digital Health Technologies, NICE adopts the definition of Innovative Medicines Initiative Get Real, data not collected in the context of RCTs, but either primary research data collected in a manner which reflects how interventions would be used in routine clinical practice or secondary research data

derived from routinely collected data (14). The definition of RWD used by the Association of the British Pharmaceutical Industry is data obtained by any non-interventional methodology (15). One of the commonly used definitions of RWD is that of the US Food and Drug Administration (FDA), "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources"(16). Another widely cited study regarding the definition of RWD is Makady et al. (1). In their study, RWD was categorised into four groups focusing more on the study design used to collect the data while the FDA definition highlights the frequency of collecting health-related information and the regularity of obtaining the information. Each definition has relatively large operational flexibility, which could potentially introduce bias into the data extraction. For instance, an observational study which collected patient data *once* from routine clinical practice can be classified differently depending on definitions. Under the definition of FDA, it is not RWD as the data is not routinely collected. On the other hand, it is RWD under the definition in Makady et al., data collected in a non-experimental setting. Requiring data to meet both definitions can help to reduce the discretionary interpretation of RWD. However, given that there is no consensus on the definition of RWD, the definition can be questioned by other researchers who have different views. Hence, this study uses two working definitions adopting two of the categories of RWD explored by Makady and his colleagues' study together matching the FDA's definition of RWD (figure 2). In this research, RWD is defined as the data relating to patient health status and/or the delivery of health care routinely collected from non-experimental settings (working definition 1) or non-interventional setting (working definition 2).

Step 1 Appraisal selection

The first step of the research identifies the NICE TA guidance which meets the eligibility criteria. TA guidance are publicly available on the NICE website (www.nice.org.uk). Relevant appraisal documents including the final scope, the manufacturer's submission, the evidence review group (ERG) report, and the final appraisal determination are available for each appraisal. The appraisal documents are reviewed to establish whether RWD is used to determine any components of the economic evaluation.

Data sources

This research exclusively includes single-technology appraisals (STA) of oncology medicines. Figure 3 shows the inclusion and exclusion criteria. One aim is to understand how and where RWD has been used in the appraisal process. Therefore, it is necessary that the appraisal process should be identical. However, the STA and multiple technology appraisal (MTA) processes differ substantially. It is challenging to gather same information in the MTA process as the MTA has different format of appraisal documents to assess several drugs or treatments used for one or more condition (17). Besides, STAs are the predominant form in practice, 93% of appraisals of oncology. The small number of the MTAs, only eighteen oncology appraisals, limits the scope for such a comparison. Therefore, this study focuses on STAs, which assess a single treatment. It also limits analysis to appraisals published between January 2011 and May 2021 in order to have a long enough time period to capture potential changes over time in how RWD has been used but also recognising that STAs from earlier years might be of less interest because enthusiasm for RWD was largely absent. Here, the date when guidance was published refers to the date of issuing the final appraisal determination document (FAD) which can be regarded as an end point of the evidence synthesis process (in the absence of a successful appeal).

Operational separation

Following the inclusion and exclusion criteria, appraisals are identified. Among these appraisals, some TAs have more than one clinical indication or combination therapy. It is possible that different evidence was used for the different patient populations in the appraisal. Hence, these appraisals are separated by clinical conditions or treatment lines and reviewed in order to avoid losing information. For example, olaparib for maintenance treatment of relapsed platinum-sensitive ovarian, fallopian tube or peritoneal cancer (NICE TA620) has two separate recommendations for different conditions. While a patient who has a BRCA1 or BRCA2 mutation and has had three or more courses of platinum-based chemotherapy is eligible for the treatment, a patient who has a BRCA1 or BRCA2 mutation and has had two courses of platinum-based chemotherapy is able to use the treatment within Cancer Drug Fund. Consequently, these indications are included separately in the analysis.

Step 2: Data extraction

A detailed protocol is developed to guide the extraction of essential data for each appraisal in order to investigate the use of RWD in NICE technology appraisals in a systematic and reproducible manner. The protocol is designed to extract information from both the manufacturer's submission and the final appraisal document regarding where RWD was used in the manufacturer's cost-effectiveness analysis, and to determine the extent to which the committee supported the use of RWD in these appraisals and understand what factors are associated with supporting or not supporting the use. **Error! Reference source not**

found.Error! Reference source not found. shows the structure of the data extraction template. In summary, the extraction tool consists of three parts – general information, explanatory variables, and outcome variables. The outcome of interest being the use of RWD. The outcome variables record use or non-use of RWD for different elements of the economic evaluation. The tool extensively includes important elements of an economic evaluation. The study will analyse the data to investigate patterns in the use of RWD in NICE appraisals, and the association between several factors and the use of RWD. Explanatory variables are suggested based on the hypotheses presented under Step 4: data analysis. All items in the extraction template and how to code them are described in the glossary (Supplement 1). To convey the type of information to be extracted, some examples from a preparatory review are presented in the glossary.

Parametric and non-parametric use

This protocol distinguishes two categories of outcome variable, parametric and non-parametric use of RWD. Parametric use of RWD is the use of such data to define the numerical value of a specific variable in the economic evaluation, whereas non-parametric use is where data are utilised to develop the model structure or to determine the scope of the evaluation. For example, when RWD are used to estimate survival, this will be counted as parametric use with respect to clinical outcome (OS/PFS). Parametric use is reviewed and recorded for the intervention and comparators separately as different data could be used in the cost-effectiveness analysis. An example of non-parametric use of RWD can be found in the appraisal of palbociclib for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495). In this appraisal, the company

used information from a study of medical records to determine the subsequent treatments to be assumed in the economic model. This case is regarded as non-parametric use since RWD was used to specify the treatment sequence but not the quantity and cost of subsequent treatment.

Parametric and non-parametric use of RWD and the different categories shown in figure 4, facilitate more consistent data extraction by highlighting the different ways RWD might be used, and provide greater flexibility to the testing of hypotheses regarding use of RWD, and the exploration of ways to measure the intensity of use of RWD.

Coding

A key issue with respect to improving the reliability of data extraction is how many distinct variables are identified and how finely divided are the potential responses to the variables in the extraction template. There are two options, in order not to lose information, have many distinct variables with binary responses, or merge many variables but have multi-level responses. In case of the variables coarsely divided, the outcome of the extraction is so blunt that it cannot fully capture how RWD is used. Likewise, the variables overscrupulously divided are less likely to provide valid outcome to show the pattern of the use of RWD in the analysis. It is closely linked to how far should it break down the levels of each coding. In an effort to sophisticatedly divide variables, the template took an “including all and combining trivia” approach. It helps to include all relevant variables where potentially data can be used, but also to list variables more concisely by merging unnecessarily trivial variables so that the outcome of the extraction can be concretely analysed. Based on two categories, the parametric and non-parametric use of RWD, the areas where data are likely to be used are carefully searched. As a backbone of the extraction structure, dividing two categories helped

to search the component systematically. Under the parametric use, the clinical effectiveness, health utility, cost side were thoroughly reviewed. After sorting variables, they were aggregated if the information is minor and can be categorised into one variable. The areas where aggregation is mostly required are resource use in economic evaluation. In order to reflect the current practice, especially cost part has naturally incorporated RWD into the analysis. It is less informative to collapse the variables in resource use since the use of RWD in cost part is already expected. Also, different health technology can include different resource use reflecting its characteristics. Counting every cost part is not an accurate way to understand why and how RWD was used. Hence, the variable, resource use is not differentiated. However, variables such as volume of treatment or dose adjustment are separated from resource use as these data have potentially critical impacts on the result of economic evaluation.

Step 3: Data validation

The data extraction tool will be validated by a second researcher independently repeating the data extraction for a random sample of appraisals (20% of all appraisals). This validation is required to check the replicability of the data extraction and the clarity of the extraction tool. Any disagreements between the researchers will be resolved by discussion. Peer discussion following the validation process is important not only to check the clarity of this protocol but also to investigate any deviations caused by unclear information. It will help pinpoint where a higher degree of subjectivity may arise in the data extraction.

Step 4: Data analysis

The extracted data will be analysed quantitatively in two different ways. First, a descriptive

analysis will summarise where and how RWD has been used in appraisals. This will be supplemented by an analysis of the intensity of use of RWD in order to explore changes in the pattern of use of RWD over time and differences with respect to cancer type. Secondly, a regression analysis will be performed to investigate which factors are associated with the greater use of RWD in a company's submission. A literature review and a pilot study were conducted to identify factors potentially associated with the use of RWD. Five factors were identified and formulated into hypotheses about increased use of RWD (**Error! Reference source not found.**).

Methodological issues

The design of this data extraction protocol, in which information is reliably and repeatedly extracted across appraisals, will allow us to review evidence for the use of RWD more systematically than could be obtained from a number of case studies. However, several methodological challenges can be anticipated. This section addresses these challenges and how they might be mitigated.

Issue 1: Unclearly stated information

Overall, NICE appraisals clearly describe the data used in the evidence synthesis. However, in some cases, a result of systematic research is listed without clarification that the study is included in evidence synthesis. In appraisals, systematic literature review is carried out to identify all relevant evidence. Clinical effectiveness evidence is carefully examined and described in detail, with clear reasons for the inclusion and exclusion of studies. On the other hand, the systematic search for resource use and cost information usually enumerates

miscellaneous studies with bibliographic information and a summary, but the critical review of minor components of health cost is sometimes missing. While a manufacturer provides the result of the assessment, it is possible that some manufacturers' submissions do not clearly state whether a particular study is used to determine which elements of resource use made up the health state costs. This leaves room for discretion how to record the information.

Issue 2: Level of aggregation

An important question is the most appropriate level of aggregation. This is best illustrated with respect to healthcare costs. It would be possible to have a variable indicating use or non-use of RWD for every single element of cost (distinguishing GP visits, frequency of hospitalization, and so on). At the opposite extreme there could be a single cost variable which indicated whether RWD was used for any element of cost. The more aggregated the measure the greater the loss of information, but some elements of cost are much more important than others and the potential analyses of the use of RWD will multiply greatly if there is no attempt at aggregation. The current protocol tries to balance the advantages and disadvantages of different levels of aggregation by combining several elements into a health state cost variable but distinguishing other important components of cost such as, such as volume of treatment, dose adjustment and resource use for adverse events.

Issue 3: No consensus on the definition of RWD.

This research uses the definition of RWD by FDA. The distinctive part of the definition used in this research is 'routinely collected' data from 'non-experimental study'. Although the definition provides specific and clear definition for this research, there is no consensus as how best to define RWD. It is likely that other definitions of RWD are preferred by other

researchers and the data extracted will be influenced by the definition of RWD chosen. Two definitions of RWD are used in this protocol. Although the definition can be questioned by other researchers who have different views, the various definitions overlap considerably. It is thus unlikely there will be a marked divergence in the data extracted when using the different definitions.

Design to mitigate methodological issues

Several operational rules are designed to minimise bias likely to come from the methodological issues encountered in the data extraction. First, the study will record the unclear information as 'no RWD.' The separation of 'not clear' is an intuitive way to extract the data, however, it is not useful for the analysis. The code 'not clear' cannot be independently analysed. It will be combined into 'no RWD' when analysing the data. In addition, having a 'not clear' category is unlikely to improve data quality. It is also closely linked to the reason for using binary code in this research. Decomposing levels of codes into several small parts can help to extract the information as it is. However, it is more likely to increase the complexity since trivial information is individually recorded. The extracted trivial data should be interpreted based on another operational rule. It is subject to increased error, particularly when testing hypotheses. For these reasons, the benefit of breaking up the level of codes into multilevel does not outweigh the benefit of binary codes while separation is much more time consuming. Instead of adapting multilevel codes, this study suggests an alternative, an intensity analysis which helps to understand valid difference within diverse patterns of the use of RWD. When looking at the pattern of the use of RWD, the intensity of the use of RWD will be analysed. Simply counting the number of times RWD is used is not an accurate way to understand why and how RWD was used. Alternatively, this study focuses on

variables which are more critically reviewed in appraisal. Variables such as survival outcome, volume of treatment and choice of comparators are more likely to influence estimated cost-effectiveness. Especially, the survival outcome is the most important information in both clinical and cost-effectiveness as well as one of the controversial areas where to use RWD. The intensity analysis is a framework to show whether RWD is used in these components alongside the quantity of the use of RWD. It can offer more benefits in deeper understanding of the use of RWD than counting all miscellaneous uses of RWD.

STRENGTHS AND LIMITATIONS

To the best of the authors' knowledge, this is the first study protocol to investigate to what extent RWD has been used in NICE appraisals. It allows the practice of extracting information to be reproducible, systematic and transparent. Strengthening the reproducibility and transparency of extracting process can maximise the understanding of the use of RWD by allowing more accurate interpretation and use of their findings. This protocol could be relevant to researchers or HTA agencies who aim to understand how various data resource was used in HTA in a context of England. The findings of this protocol can provide the full picture of the use of RWD in NICE appraisal over ten years. Moreover, the study findings could add more value under the discussion of NICE which tries to broaden the evidence in NICE practice.

The protocol has the limitation that it has been developed to study the use of RWD in NICE appraisals of oncology drugs. Consequently, the data extraction protocol may not be fully applicable to appraisals in other disease areas or to the different practice of other HTA bodies.

Since the documentation is significantly different depending on each country's context, it is not feasible for the protocol to extract the same information as in English context. However, many of the distinctions are of wider application, e.g. parametric vs non-parametric use of RWD, and the taxonomy of where in an economic evaluation it might be relevant to look for use of RWD. Also, the hypotheses are potentially of wider application. The results are going to be specific to NICE but otherwise the structure of this research has wider application. Although not fully transferrable, this protocol can be modified for use in other HTA contexts.

PATIENT AND PUBLIC INVOLVEMENT

No patient involved.

ACKNOWLEGEDMENTS

The authors thank to David Lugo-Palacios for useful advice on clarity of data extraction and validation process.

CONTRIBUTORS

Both authors contributed to conceptualising and designing the study. JK drafted the protocol manuscript. JC revised the manuscript for important intellectual content and contributed to the methodology.

FUNDING STATEMENT

JK is supported by the Centre for Cancer Biomarkers, University of Bergen funded by the Research Council of Norway grant number (223250). The funder was not involved in the design of the protocol or decision to submit the protocol for publication, nor will they be involved in any aspect of the study conduct.

COMEPTING INTEREST STATEMENT

None declared.

FULL REFERENCES

1. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. Value Heal [Internet]. 2017;20(7):858–65. Available from: http://dx.doi.org/10.1016/j.jval.2017.03.008

2. Chan K, Nam S, Evans B, Deoliveira C, Chambers A, Gavura S, et al. Developing a framework to incorporate real-world evidence in cancer drug funding decisions: The Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration. BMJ Open. 2020 Jan 7;10(1).

3. Latimer N. Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data - PubMed - NCBI. Med Decis Mak [Internet]. 2013 [cited 2019 May 14];33(6):743–54. Available from:

https://www.ncbi.nlm.nih.gov/pubmed/27905716

4. Davies C, Briggs A, Lorgelly P, Garellick G, Malchau H. The “hazards” of extrapolating survival curves. Med Decis Mak [Internet]. 2013 Apr [cited 2020 Aug 17];33(3):369–80. Available from: https://pubmed.ncbi.nlm.nih.gov/23457025/

5. Katkade VB, Sanders KN, Zou KH. Real world data: An opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. J Multidiscip Healthc. 2018;11:295–304.

6. Littlejohns P, Chalkidou K, Culyer AJ, Weale A, Rid A, Kieslich K, et al. National Institute for Health and Care Excellence, social values and healthcare priority setting [Internet]. Vol. 112, Journal of the Royal Society of Medicine. SAGE Publications Ltd; 2019 [cited 2021 May 11]. p. 173–9. Available from: https://journals.sagepub.com/doi/full/10.1177/0141076819842846

7. Schlander M. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Springer New York; 2007.

8. Broader types of data to be used in development of NICE guidance | News and features | News | NICE [Internet]. [cited 2020 May 5]. Available from: https://www.nice.org.uk/news/article/broader-types-of-data-to-be-used-in-development-of-nice-guidance

9. Wang S, Goring SM, Lozano-Ortega G. Inclusion of real-world evidence in submission packages to health technology assessment bodies: What do current guidelines indicate? Value Heal [Internet]. 2016 May 1 [cited 2020 Jun 22];19(3):A287. Available

from: http://www.valueinhealthjournal.com/article/S1098301516008317/fulltext

10. Pietri G, Masoura P. Market Access and Reimbursement: The Increasing Role of Real-World Evidence. Value Heal [Internet]. 2014 Nov 1 [cited 2020 Jun 22];17(7):A450–1. Available from: http://www.valueinhealthjournal.com/article/S1098301514031465/fulltext

11. Harwood M, Deighton A, Mickle A, Qian C, Szabo S. The use of real-world data in health technology assessment of medications for rare diseases. Value Heal [Internet]. 2019 Nov 1 [cited 2020 Sep 10];22:S863. Available from: http://www.valueinhealthjournal.com/article/S1098301519348296/fulltext

12. Roberts MH, Ferguson GT. Real-World Evidence: Bridging Gaps in Evidence to Guide Payer Decisions. PharmacoEconomics - Open [Internet]. 2021 Mar 18 [cited 2021 Jun 29];5(1):3–11. Available from: https://doi.org/10.1007/s41669-020-00221-y

13. Bullement A, Podkonjak T, Robinson MJ, Benson E, Selby R, Hatswell AJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. Int J Technol Assess Health Care [Internet]. 2020 Aug 1 [cited 2021 May 12];36(4):388–94. Available from: https://doi.org/10.1017/S0266462320000434

14. National Institute for Health and Care Excellence (NICE). Evidence Standards Framework for Digital Health Technologies Contents. 2019.

15. Association of the British Pharmaceutical Industry. Demonstrating value with Real World data | ABPI [Internet]. 2011 [cited 2021 Jun 30]. Available from: https://www.abpi.org.uk/publications/real-world-data/

16. FDA. Real-World Evidence [Internet]. 2020 [cited 2020 May 6]. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

17. Guide to the multiple technology appraisal process [Internet]. [cited 2020 Aug 7]. Available from: www.nice.org.uk

Figure Legends

Figure 1 The process of the study

Figure 2 Flowchart to decide the definition of RWD

Figure 3 Inclusion/exclusion criteria

Figure 4 The framework for data extraction

Figure 5 Hypotheses about increased use of RWD

Supplement Legends

Supplement 1 Glossary of variables in extraction template

Figure 1 The process of the study

84x11mm (240 x 240 DPI)

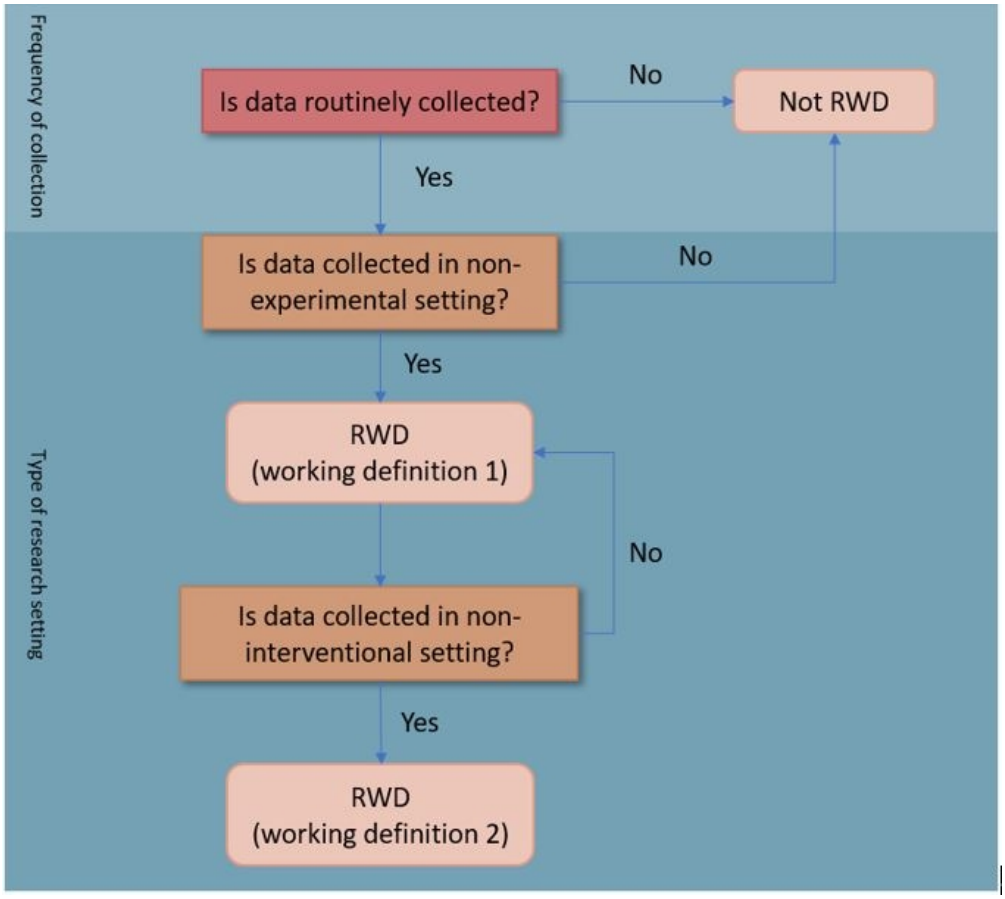


Figure 2 Flowchart to decide the definition of RWD

71x64mm (240 x 240 DPI)

Inclusion criteria
- STA of oncology medicine
- Appraisals issued from January 2011 to May 2021
Exclusion criteria
- Appraisal of technology for preventing the complications of cancer
- Appraisal of surgical practice and other therapeutic therapies
- Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals)

Figure 3 Inclusion/exclusion criteria

131x35mm (240 x 240 DPI)

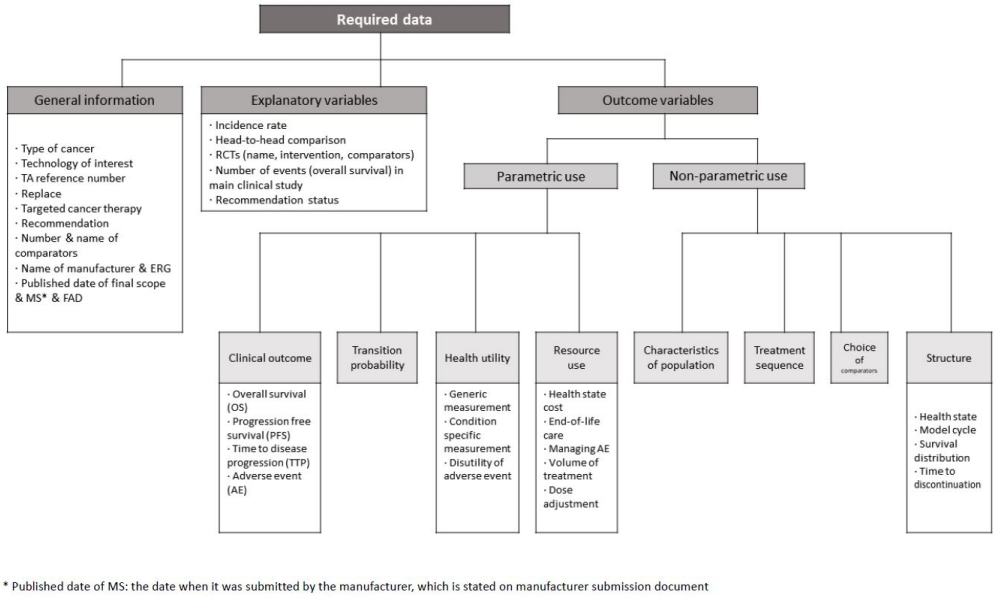


Figure 4 The framework for data extraction

141x84mm (240 x 240 DPI)

1) Poor internal/external validity of the clinical trial is associated with greater use of RWD.
2) Absence of direct (head-to-head) comparison is associated with greater use of RWD.
3) Low incidence rate of the disease is associated with greater use of RWD.
4) Immature survival data in the clinical trial are associated with greater use of RWD.
5) The technology having been recommended in previous NICE TA guidance is associated with greater use of RWD.

Figure 5 Hypotheses about increased use of RWD

93x24mm (240 x 240 DPI)

Supplement 1 Glossary of variables in extraction template

General information		
Variable	**Explanation**	**Coding**
Type of cancer	The NICE classification of the cancer (website: https://www.nice.org.uk/guidance/conditions-and-diseases/cancer)	Bladder cancer=1, Blood and bone marrow cancer =2, Breast cancer=3, Colorectal=4, Neuroblastoma=5, Head and neck=6, Liver=7, Lung=8, Oesophageal=9, Ovarian=10, Pancreatic=11, Prostate=12, Renal=13, Skin=14, Stomach=15, Sarcoma=16
Technology of interest	The name of drug in the current appraisal. If it is combination therapy, the key technology which manufacturer focuses on will be taken here.	Narrative description
Indication	Clinical indications which are addressed in Final Appraisal Determination (FAD) document	Narrative description
TA number	the reference number of the technology guidance	Narrative description
Replaced	Whether TA guidance has been replaced or not	None= 0 If replaced, the TA reference number and year
Targeted cancer therapy	Treatment that uses drugs or other substances to identify and attack specific types of cancer cells	Non-targeted therapy = 0, targeted therapy = 1, not sure = Narrative description
Recommendation	the classification of recommendations made by the NICE committee in FAD document - Not recommended: 0 - Recommended (in line with marketing authorisation): 1 - Recommended (in line with marketing authorisation) in CDF:2 - Optimised: 3 - Optimised in CDF: 4	Not recommended=0, recommended=1, recommended (cdf)=2, optimised=3, optimised (cdf)=4
number of comparators	Count the number of comparators in each manufacturer submission or FAD document. The information in manufacturer submission and FAD is recorded in the separated rows (manufacturer row/committee row).	Number in the manufacturer's submission
name of comparators	Record the name of comparators in manufacturer submission or FAD document	Narrative description

name of manufacturer	the name of manufacturer in manufacturer submission	Narrative description
name of the ERG	the name of the ERG (evidence review group)/AG (assessment group) in ERG critiques or AG reports	Narrative description
published date of final scope	the date of final scope as MM/YYYY	Date (MM/YYYY)
published date of manufacturer	the date of manufacturer submission as MM/YYYY.	Date (MM/YYYY)
published date of FAD guidance	the date of FAD document as MM/YYYY	Date (MM/YYYY)
Explanatory variables		
Variable	**Explanation**	**Coding**
Incidence (rate, year)	The rate would be recorded as it is in the appraisal. Incidence rate could be found in the final scope document or in manufacturer submission document. If the figures are not identical in each document, the latest rate is recorded.	Narrative description
H2H	Whether the head-to-head clinical trial of a technology of interest exists or not, which compares with agreed comparators. The information is most likely to be found in the section: Identification and selection of relevant studies in clinical effectiveness part.	no=0, yes=1, yes but some comparators missing =2
• ITC	ITC (indirect treatment comparison). The information could be found in the section: Indirect and mixed treatment comparisons in clinical effectiveness part.	no=0, yes=1
• RCT (technology of interest)	Main RCT used in the appraisal: the name of the H2H RCT, if it exists. Unless there is an H2H, RCT refers to the clinical trial of technology of interest in the ITC.	no=0, yes=1
- Name of RCT	The name of the aforementioned RCT	Narrative description
- Intervention in RCT	The name of the intervention used in the aforementioned RCT. This variable helps to identify the main technology in RCT when technology is appraised as combination therapy.	Narrative description
- Comparators in RCT	The comparator of the aforementioned RCT	Narrative description
- Size of RCT	The number of participants in the aforementioned RCT	Number
- Median duration of follow-up	The median duration of follow-up in the aforementioned RCT. If it is not reported, record as NR (not reported).	Unit: month Not reported = ..
• Anchored/unanchored	"Anchored" means that RCT of technology of interest exists, and the RCT has been linked to any other studies which evaluate the drug's effectiveness.	Not anchored=0, Anchored =1

	"Unanchored" means that the clinical outcome study doesn't have any comparators which connect to other studies. For example, comparing a single-arm study with a single-arm study is "unanchored". Also, RCTs compared without common comparators in ITC is "unanchored".	
• MAIC/STC	Matching adjusted indirect comparison (MAIC), Simulated Treatment Comparison (STC). A methodology of making adjustment to increase the comparability of two distinct populations mostly among unanchored studies. But it could be used in anchored studies in case where the two populations in ITC is starkly different from each other.	Naive=0, MAIC=1 STC=2 Other methods=3
Risk of bias (RoB) of RCT (direct quotation)	In order to evaluate the internal validity of RCTs, the risk of bias, which was reported in the ERG report, will be recorded here. Information is available at the quality assessment part of the ERG report. The ERG statement is directly quoted.	Direct quotation from ERG documents
• Risk of bias in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into four groups following the number of risk factors.	High/good quality without mentioned weakness= 0, risk factor 1 (low) =1, risk factor 2-3 (moderate)=2, risk factor 4 (high) =3
External validity of RCT	As narrative accounts, generalisability of RCT is reported in the ERG report whether the population of RCT properly represents the UK general population in terms of aging structure, health status and health care practice (practice-dose, subsequent treatment, etc.).	Direct quotation from ERG documents
• External validity in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into three groups following the severity of generalisability assess by ERG.	Representative without mentioned weakness= 0, Representative but minor concerns =1, Questionable generalisability =2
Previously recommended in other indication	Whether the technology has been recommended for other types of cancers besides the current indication of the technology.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in other indication	If it was recommended for other indications, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Previous recommended treatment in the same cancer	Whether the technology has been recommended for other treatment lines in the same type of cancer.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in the same cancer	If it was recommended for other treatment lines in the same cancer category, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Maturity of survival data in clinical trial	The data maturity is examined by looking at the number of events (deaths) of intervention arm in clinical trials.	Direct quote from manufacturer submission
• Maturity (grade)	The direct quotation will be classified into three groups following the data cut point, 20% and 50 % of the number of events.	Immature (number of events < 20%) =1, Relatively immature (20%≤number of

			events≤50%)=2 Mature (number of events < 50%) =3
Outcome variables			
Variable	**Explanation**	**Coding**	**Example**
characteristic of population	Whether RWD are used to determine the characteristic of population, including the initiation age and health performance status (ECOG) or not. - Soft use: when RWD are supplementary evidence to decide the population characteristics - Hard use: when RWD determine the characteristics of population in economic evaluation	No RWD = 0 Yes, data from RWD = 1	- Pomalidomide, in combination with low-dose dexamethasone, for treating multiple myeloma in adults at third or subsequent relapse (NICE TA427): baseline patient characteristics were obtained from RWD collected from a hospital population since the majority of the trial populations were previously untreated, which was different from target population.
treatment sequence	Whether RWD are used to determine the subsequent treatment option or not. After the disease progression onto the later stages of cancer treatments, patients are likely to receive idiosyncratic subsequent treatments. The pattern of subsequent treatment for cost-effectiveness analysis could be observed by RCT or RWD.	No RWD = 0 Yes, data from RWD = 1	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): a study of medical records was used to determine the treatment sequence.
choice of comparator	Whether RWD are used to choose the comparators in economic evaluation or not. Although comparators are chosen based on the current clinical guideline, drug utilisation data or clinical expert opinion are frequently referred to find the most relevant comparators in evaluation.	No RWD = 0 Yes, data from RWD = 1	- Ixazomib with lenalidomide and dexamethasone for treating relapsed or refractory multiple myeloma (NICE TA505): the manufacturer considered that lenalidomide was appropriate comparator based on IMS market research data (lenalidomide, 69% market share and panobinostat, 7%).
structure (health state)	Whether RWD are used to determine the health state such as stable, progression, and death in a given model. Information is available at health state in the model of cost-effectiveness analysis in manufacturer submission documents.	No RWD = 0 Yes, data from RWD = 1	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): the model health state of post-progression was specified based on a retrospective patient medical record review study.
structure (model cycle)	Whether RWD are used to determine model cycle or not. Model cycle, hereby, means that the duration between different health states, which can be influenced by the severity of conditions.	No RWD = 0 Yes, data from RWD = 1	N/A **
Structure	Whether RWD are used to decide the survival	No RWD = 0	- Larotrectinib for treating advanced solid tumours with

(survival distribution of intervention)	distribution of intervention or not. Since survival rate observed in RCTs is immature, it is necessary to extrapolate the survival rate for analysis. In order to choose proper survival distribution, the goodness of fit is tested (AIC, BIC). Also, the clinical plausibility is asked to validate the distribution. In this case, the alternative data can be utilized. - If RWD is utilised for choosing distribution, mark as "hard use". - If RWD is utilised as supplementary evidence for the chosen distribution, mark as "soft use".	Yes, data from RWD = 1	NTRK fusions (NICE TA630): UK all-cause mortality data were used to assess the clinical acceptability of distributions whether patient overall survival exceeded current UK life expectancy
Structure (survival distribution of comparator)	Whether RWD are used to decide the survival distribution of comparator or not. As survival distributions of intervention and comparators are separately determined, the extraction tool approach it independently. Apply the abovementioned description on survival distribution of intervention to comparator in this row.	No RWD = 0 Yes, data from RWD = 1	
Structure (Time to discontinuation of intervention)	Whether RWD are used to decide the time to discontinuation of intervention or not. The time to discontinuation is likely to be decided by 1) simply adopting discontinuation rule in trials, 2) formulating distribution of discontinuation, or 3) clinical experts' opinion. - If RWD are used for designating the time to discontinuation, mark as "hard use" - If RWD are used as supplementary evidence for designating the time to discontinuation, mark as "soft use". - If clinical experts' opinions are used for designating the time to discontinuation, it is not regarded as	No RWD = 0 Yes, data from RWD = 1	- Lorlatinib for previously treated ALK-positive advanced non-small-cell lung cancer (NICE TA628): The plausibility of the extrapolation of time on treatment was validated by UK RWD, hospital network data.

	RWD.		
Structure (time to discontinuation of comparator)	Whether RWD are used to decide the time to discontinuation of comparator or not. Apply the above-mentioned description on time to discontinuation of intervention to comparator in this row.	No RWD = 0 Yes, data from RWD = 1	
Clinical outcome (OS) intervention	Whether RWD give the figure for overall survival (OS) of intervention or not. In order to measure the Quality Adjusted Life-Years (QALYs), it is necessary to extrapolate overall survival based on observed data on survival. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1	- Nivolumab for adjuvant treatment of completely resected melanoma with lymph node involvement or metastatic disease (NICE TA558): the survival model applied the registry data (American Joint Committee on Cancer; AJCC) to both treatment arms after a certain time point.
Clinical outcome (PFS) intervention	Whether RWD give the figure for progression free survival (PFS) of intervention or not. The progression of disease is important for economic evaluation model in terms of health state transitions and treatment switching. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (RR) intervention	Whether RWD provides the response rate (RR) for the intervention or not. The effectiveness of cancer treatment is often shown by responses of tumour cells, which is evaluated by the RECIST criteria or other criteria. The response rate data would be collected in RCT or other type of data.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (TTP) intervention	Whether RWD give the figure for time-to-progression (TTP) of intervention or not. Some cancer treatments show their clinical effectiveness not through the progression free survival (PFS), but alternatively through time-to-progression.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (AE) intervention	Whether RWD give the figure of adverse event (AE) of intervention or not. Adverse events are crucial information for the estimation of the QALYs. The adverse events are collected in RCT. However, RWD, including cohort studies, retrospective studies, or	No RWD = 0 Yes, data from RWD = 1	- Blinatumomab for treating acute lymphoblastic leukaemia in remission with minimal residual disease activity (NICE TA589): retrospective non-interventional cohort study collected from 2000 to 2017 was used to inform the clinical outcome of comparators as well as

	other type of studies, also provide the information of adverse events, which cannot be found in RCT.		adverse event.
Clinical outcome (OS) comparators	Whether RWD give the figure of overall survival (OS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1	Refer to the variable, clinical outcome (OS) intervention
Clinical outcome (PFS) comparators	Whether RWD give the figure for the progression free survival (PFS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (RR) comparators	Whether RWD provide the response rate (RR) of comparators or not.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (TTP) comparators	Whether RWD provide the time-to-progression (TTP) of comparators or not.	No RWD = 0 Yes, data from RWD = 1	N/A**
Clinical outcome (AE) comparators	Whether RWD provide the figure adverse events (AE) for the comparators or not.	No RWD = 0 Yes, data from RWD = 1	Refer to the variable, clinical outcome (AE) intervention
Transition probability	Whether RWD provide the transition probability from one state to other state, if it is applicable.	No RWD = 0 Yes, data from RWD = 1	- Pembrolizumab for treating melanoma with high risk of recurrence (NICE TA553): electronic health records (Flatiron database) collected by cancer care providers in the US was used to model transition from the "locoregional recurrence (LR)" state to the "distant metastases" and life tables for transition from the LR to "death" state.
Health utility of health state (generic)	Whether health state utility survey of generic measurement is done in RWD or RCT. Health state utility is necessary information for the estimation of the QALYs. Generic health utility measurement, EQ-5D, is frequently used. There is national tariff of EQ-5D to get the scores. Hereby, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1	N/A**
Health utility of health state (condition-specific)	Whether health state utility survey of condition-specific measurement is done in RWD or RCT. In cancer treatment, condition-specific measurement is commonly adopted. Similar to the previous row, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1	N/A**

Disutility of adverse events	Whether survey of collecting disutility data is done in RWD or RCT. As adverse events are likely to reduce the patient's quality of life, the disutility of adverse events is included in estimates. The way of collecting survey (RWD or RCT) is drawn to attention.	No RWD = 0 Yes, data from RWD = 1	N/A**
Resource use (Health state cost) common	Whether resource use for estimating health state cost is derived from RWD or RCT. In economic evaluation, the unit cost mostly comes from the national reference cost. The total cost is calculated by the total resource use (volume of technology and health care services) multiplied by the reference cost. Here, the only resource use is focused in data extraction.	No RWD = 0 Yes, data from RWD = 1	- Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies (NICE TA559): RWD was used for estimating the cost of inpatient admission (data: Hospital Episode Statistics), the cost of home care and hospice (data: National Audit Office), and GP time (data: Personal Social Services Research Unit; PSSRU).
Resource use (end-of-life care) common	Whether resource use for estimating end-of-life care is derived from RWD or RCT. Resource use of terminal cancer patients is not frequently reported in the RCT providing the treatment effect. Therefore, other data resources, including RCTs of other technologies, provide the information of resource use in the end-of-life care.	No RWD = 0 Yes, data from RWD = 1	N/A**
Resource use (Managing AE) intervention	Whether resource use for managing adverse events of intervention is derived from RWD or RCT. Resource use of managing adverse events is reported in RCTs as well as in other types of researches which can provide alternative perspectives.	No RWD = 0 Yes, data from RWD = 1	N/A**
Resource use (volume of treatment) intervention	Whether resource use for volume of treatment of intervention is derived from RWD or RCT. In this study, scope of the volume of treatment is limited to the frequency of treatment, frequency of administration, and market share.	No RWD = 0 Yes, data from RWD = 1	- Fulvestrant for treating untreated locally advanced or metastatic oestrogen-receptor positive breast cancer (NICE TA503): a medical chart review study was used to determine the proportion of patient using subsequent treatment for cost calculation.
Resource use (Dose adjustment) intervention	Whether resource use for dose adjustment of intervention is derived from RWD or RCT. There are several reasons for adjusting dose such as adverse events (AEs). The dose of cancer treatments is calculated by BSA (body surface area). This study focuses only on BSA and dose adjustment due to AEs,	No RWD = 0 Yes, data from RWD = 1	N/A**

	because these information are commonly reported in NICE appraisals.		
Resource use (Managing AE) comparators	Whether resource use for managing adverse events of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1	N/A**
Resource use (volume of treatment) comparators	Whether resource use for volume of treatment of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1	Refer to the variable, resource use (volume of treatment) intervention
Resource use (Dose adjustment) comparators	Whether resource use for dose adjustment of comparators is derived from RWD or RCT. Since the intervention is a novel technology, RCTs provide less information on the adjustment. RWD could be utilised to provide more relevant information regarding dose adjustment of existing technologies which have been used in routine clinical practice.	No RWD = 0 Yes, data from RWD = 1	N/A**
* In order to detect the use of RWD in sensitivity analysis, the parametric part is duplicated. ** As data extraction is not conducted, all of examples are not available at this stage. In this case, it marked as N/A.			

BMJ Open

A protocol for data extraction: how real-world data have been used in NICE appraisals of cancer therapy

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055985.R1
Article Type:	Protocol
Date Submitted by the Author:	04-Nov-2021
Complete List of Authors:	Kang, Jiyeon; London School of Hygiene and Tropical Medicine Faculty of Public Health and Policy, Department of Health Services Research and Policy; Cairns, John; London School of Hygiene and Tropical Medicine, Department of Health Services Research and Policy
Primary Subject Heading:	Health economics
Secondary Subject Heading:	Health economics
Keywords:	ONCOLOGY, HEALTH ECONOMICS, HEALTH SERVICES ADMINISTRATION & MANAGEMENT

A protocol for data extraction: how real-world data have been used in NICE appraisals of cancer therapy

Jiyeon Kang[1], John Cairns[1]

[1]Department of Health Service Research and Policy,

Faculty of Public Health and Policy,

London School of Hygiene and Tropical Medicine, London, UK

Corresponding author: Jiyeon Kang (Jiyeon.Kang@lshtm.ac.uk)

15-17 Tavistock place, London, WC1H 9SH, UK

Word count: 4,603

Keywords: Real-world data (RWD), Health Technology assessment (HTA), economic evaluation, oncology medicine, National Institute for Health and Care Excellence (NICE)

ABSTRACT

Introduction

Due to the limitations of relying on randomised controlled trials, the potential benefits of real-world data (RWD) in enriching evidence for health technology assessment (HTA) are highlighted. Despite increased interest in RWD, there is limited systematic research investigating how RWD has been used in HTA. The main purpose of this protocol is to extract relevant data from National Institute for Health and Care Excellence (NICE) appraisals in a transparent and reproducible manner in order to determine how NICE has incorporated a broader range of evidence in the appraisal of oncology medicines.

Methods and analysis

The appraisals issued between January 2011 to May 2021 are included following inclusion criteria. The data extraction tool newly developed for this research includes the critical components of economic evaluation. The information is extracted from identified appraisals in accordance with extraction rules. The data extraction tool will be validated by a second researcher independently. The extracted data will be analysed quantitatively to investigate to what extent RWD has been used in appraisals. This is the first protocol to enable data to be extracted comprehensively and systematically in order to review the use of RWD.

Ethics and dissemination

This study is approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine on 14 November 2019 (17315). Results will be published in peer-reviewed journals.

STRENGTHS AND LIMITATIONS OF THIS STUDY

Strengths

- This protocol enables data to be extracted in a transparent and systematic manner for the study of how RWD has been used in NICE appraisals.

- It includes all the different ways an economic evaluation might use RWD.

- This study facilitates systematic understanding of the use of RWD in NICE appraisals over the last 10 years.

Limitations

- Since it is focussed on cancer, the methods and eventually the findings are to some extent cancer-specific.

- The protocol could be modified to reflect the HTA context in different countries although the extraction protocol is not fully applicable to the practice of other HTA bodies as much of the protocol reflects the NICE appraisal process.

INTRODUCTION

In the last few years, interest in real-world data (RWD) has grown in health care decision-making (1). Health Technology Assessment (HTA) refers to the systematic evaluation of clinical- and cost-effectiveness of health technology (2,3). It requires valid and reliable information for the evaluation. Randomised controlled trials (RCTs) have mainly provided the information (4). However, it is challenging to meet all information needs from RCTs since the new generation of therapies pose several assessment challenges. For example, when treatment options are expanding rapidly, it is increasingly unlikely that there are RCTs featuring all the relevant comparators. Furthermore, the traditional design of RCTs is possibly less appropriate for new technologies such as those targeting rare genetic mutations that it is difficult to obtain enough size of clinically relevant study population (5). Moreover, RCTs tend to include strictly controlled populations. Restricted population makes replication of finding challenging (6). Another barrier to obtaining all the relevant information for HTA from RCTs concerns the extrapolation of survival. RCTs observe the clinical outcome for a certain period. Extrapolation is required in order to incorporate the survival data from RCTs in the health economic model (7). It is more challenging to identify the most appropriate

extrapolation the shorter the duration of the trial. If survival data from RCTs are based on a very limited observation period, the extrapolation of the survival curve is likely to fail to predict the long-term effect (8).

The potential benefits of RWD in enriching evidence for HTA are highlighted by the limitations of relying on RCTs (9). This research focuses on the use of RWD in HTA by the National Institute for Health and Care Excellence (NICE). NICE has achieved an international reputation for rigorous development and application of scientific methods to appraise new health technologies to provide its decisions with robust and fair justification (10). More importantly, NICE is noted for the transparency of its processes, responsiveness to change, and commitment to using the best available evidence (11). The evidence is structurally well-documented enough to find the key information and is available on the NICE website. Therefore, review of these appraisals can provide comprehensive information on the evidence used for decision-making. In April 2020, NICE signalled its intention to integrate broader types of data in developing NICE guidance (12). Although it is primarily a statement of intent, it is not a new development in NICE practice since NICE already incorporates a diverse range of published scientific evidence when developing its guidance on health technologies. For example, UK audit data (TA255, 2012), Hospital Episode Statistics (TA559, 2018) and registry data such as the Edinburgh Ovarian Cancer Database (TA598, 2019), Surveillance, Epidemiology, and End Result program (TA562, 2019) have been used in the development of NICE technology appraisal (TA) guidance. While a wide range of data are already used in NICE guidance, there is limited understanding regarding how and where RWD has been used, and in which circumstances RWD is accepted as relevant. Research is required to investigate systematically patterns in the use of RWD and to understand the driving forces

behind its use in NICE appraisals.

Several researchers have reviewed practice across HTA bodies (13,14) or reported the use of RWD in HTA (15). However, little systematic research has been conducted. Important information is missing such as how they included literatures without selection bias, which parts of evidence were reviewed, whether they have clearly defined RWD and justified or explained why this definition is relevant and how different HTA systems were compared given their different practices. Roberts et al. addressed the potential role of RWD in bridging the evidence gaps (16). However, they illustrate the use of RWD with a few examples, rather than providing a fuller picture of current practice when using RWD. Bullement et al. recently reviewed how RWD informed single technology appraisals of cancer drugs in NICE (17). Although this study follows a more systematic approach to the review of the use of RWD, a data extraction table was not provided and the authors focused only on how RWE influenced the cost-effectiveness analysis, and not on how RWE was used to support or establish the appraisal. Due to limited information presented concerning the review process in this study, it is unclear whether the information presented provides a full picture of the use of RWD. Bullement et al. included 113 STAs issued between April 2011 and October 2018. As interest in RWD is increasing over time, it may miss relevant information from recent years. Therefore, the careful development of this extraction protocol is required to help extract the data systematically from appraisals, increase the reliability of the results of the analysis and permit a more detailed description of the use of RWD and analysis of factors influencing its use.

A protocol is required to ensure the consistency of data extraction so that the risk of unsystematic data collection is reduced. The main purpose of this protocol is to extract data from NICE appraisals in a transparent and reproducible manner to answer, “how has NICE

incorporated a broad range of evidence in the appraisal of oncology medicines." Without proper justification and operational rules, the data are likely to be collected erroneously, with a risk of biasing the analysis. The extracted data are expected to be objective and less biased. With such data, the analysis can provide more robust answers to questions regarding how RWD has been used in NICE technology appraisals. Furthermore, this protocol facilitates the development of a rich dataset which can highlight not just where RWD has been used but also what types of evidence have been used in the HTA process in line with NICE's interest in incorporating a broad range of evidence. The data can be analysed to answer the research questions including "how has RWD been used in NICE appraisals" and "which factors are associated with increased likelihood of the use of RWD" in depth.

METHODS AND ANALYSIS

NICE appraisal documents are identified following inclusion criteria (figure 1). The information is extracted from identified appraisals in accordance with extraction rules. The detailed extraction rules can be found in supplement 1. The extraction tool includes evidence-related information such as characteristics of the main clinical evidence and the economic evaluation model and other information. Following the tool, the information about in which part of cost-effectiveness analysis RWD was used is collected. Intensity analysis of the use of RWD and regression analysis are planned. The data extraction is planned from January 2020 to October 2021.

Definition of RWD

Before extracting information about the use of RWD in NICE technology appraisals, a definition of RWD is clearly required. RWD is an umbrella term which covers broad categories of data. Although RWD is increasingly addressed in the literature, there is no consensus over the definition. One of the commonly used definitions of RWD is that of the US Food and Drug Administration (FDA)(18). Another widely cited study regarding the definition of RWD is Makady et al. (1). Each definition has relatively large operational flexibility to be used for data extraction. For example, companies sometimes present phase 1 clinical trial as RWD. However, these data hardly provide insights in the discussion of the use of RWD in HTA. One of the plausible ways to include right data for the relevant review is to restrict the range of RWD by merging two definitions from FDA and Makady et al. Requiring data to meet both definitions can help to reduce the discretionary interpretation of RWD. Hence, this study uses a definition combining a category of the study designs of collecting RWD explored by Makady and his colleagues' study and the FDA's definition of RWD focusing on routinely collected data. In this research, RWD is defined as the data relating to patient health status and/or the delivery of health care routinely collected from non-experimental settings.

Step 1 Appraisal selection

The first step of the research identifies the NICE TA guidance which meets the eligibility criteria. TA guidance are publicly available on the NICE website (www.nice.org.uk). Appraisal documents are available for each appraisal. Among the documents, this study only reviews four type of appraisal documents, the final scope, the manufacturer's submission, the evidence review group (ERG) report, and the final appraisal determination. These documents are reviewed to establish whether RWD is used to determine any components of the

economic evaluation.

Data sources

This research exclusively includes single-technology appraisals (STA) of oncology medicines. Figure 1 shows the inclusion and exclusion criteria. One aim is to understand how and where RWD has been used in the appraisal process. Therefore, it is necessary that the appraisal process should be identical. However, the STA and multiple technology appraisal (MTA) processes differ substantially. The MTA has different format of appraisal documents to assess several drugs or treatments used for one or more condition. It is challenging to gather same information in the MTA process as different actors have the principal responsibility for producing the main evidence in each process (19). Besides, STAs are the predominant form in practice, 93% of appraisals of oncology. The small number of the MTAs, only eighteen oncology appraisals, limits the scope for such a comparison. Therefore, this study focuses on STAs, which assess a single treatment. It also limits analysis to appraisals published between January 2011 and May 2021 in order to have a long enough time period to capture potential changes over time in how RWD has been used but also recognising that STAs from earlier years might be of less interest because enthusiasm for RWD was largely absent. Here, the date when guidance was published refers to the date of issuing the final appraisal determination document (FAD) which can be regarded as an end point of the evidence synthesis process (in the absence of a successful appeal).

Operational separation

Following the inclusion and exclusion criteria, appraisals are identified. Among these

appraisals, some TAs have more than one clinical indication or combination therapy. It is possible that different evidence was used for the different patient populations in the appraisal. Hence, these appraisals are separated by clinical conditions or treatment lines and reviewed in order to avoid losing information. For example, olaparib for maintenance treatment of relapsed platinum-sensitive ovarian, fallopian tube or peritoneal cancer (NICE TA620) has two separate recommendations for different conditions. While a patient who has a BRCA1 or BRCA2 mutation and has had three or more courses of platinum-based chemotherapy is eligible for the treatment, a patient who has a BRCA1 or BRCA2 mutation and has had two courses of platinum-based chemotherapy is able to use the treatment within Cancer Drug Fund. Consequently, these indications are included separately in the analysis.

Step 2: Data extraction

A detailed protocol is developed to guide the extraction of essential data for each appraisal in order to investigate the use of RWD in NICE technology appraisals in a systematic and reproducible manner. The protocol is designed to extract information from both the manufacturer's submission (manufacturer's cost-effectiveness analysis) and the final appraisal document (the model preferred by the committee) regarding where RWD was used, and to determine the extent to which the committee supported the use of RWD in these appraisals and understand what factors are associated with supporting or not supporting the use. Figure 2 shows the structure of the data extraction template. In summary, the extraction tool consists of three parts – general information, explanatory variables, and outcome variables. The outcome of interest being the use of RWD. The outcome variables record use or non-use of RWD for different elements of the economic evaluation. The information in

base-case analysis and sensitivity analysis will separately extracted. The tool extensively includes important elements of an economic evaluation. The study will analyse the data to investigate patterns in the use of RWD in NICE appraisals, and the association between several factors and the use of RWD. Explanatory variables are suggested based on the hypotheses presented under Step 4: data analysis. All items in the extraction template and how to code them are described in the glossary (supplement 1). To convey the type of information to be extracted, some examples from a preparatory review are presented in the glossary.

Parametric and non-parametric use

This protocol distinguishes two categories of outcome variable, parametric and non-parametric use of RWD. Parametric use of RWD is the use of such data to define the numerical value of a specific variable in the economic evaluation, whereas non-parametric use is where data are utilised to develop the model structure or to determine the scope of the evaluation. For example, when RWD are used to estimate survival, this will be counted as parametric use with respect to clinical outcome (OS/PFS). Parametric use is reviewed and recorded for the intervention and comparators separately as different data could be used in the cost-effectiveness analysis. An example of non-parametric use of RWD can be found in the appraisal of palbociclib for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495). In this appraisal, the company used information from a study of medical records to determine the subsequent treatments to be assumed in the economic model. This case is regarded as non-parametric use since RWD was used to specify the treatment sequence but not the quantity and cost of subsequent treatment.

Parametric and non-parametric use of RWD and the different categories shown in figure 2, facilitate more consistent data extraction by highlighting the different ways RWD might be used, and provide greater flexibility to the testing of hypotheses regarding use of RWD, and the exploration of ways to measure the intensity of use of RWD.

Coding

A key issue with respect to improving the reliability of data extraction is how many distinct variables are identified and how finely divided are the potential responses to the variables in the extraction template. There are two options, in order not to lose information, have many distinct variables with binary responses, or merge many variables but have multi-level responses. This coding system has advantages which include avoiding information loss, and also grouping 'similar' information used during appraisals to establish patterns of the use of RWD. It is closely linked to how far should it break down the levels of each coding. In an effort to sophisticatedly divide variables, the template took an "including all and combining trivia" approach. It helps to include all relevant variables where potentially data can be used, but also to list variables more concisely by merging unnecessarily trivial variables so that the outcome of the extraction can be concretely analysed. Based on two categories, the parametric and non-parametric use of RWD, the areas where data are likely to be used are carefully searched. As a backbone of the extraction structure, dividing two categories helped to search the component systematically. Under parametric use, clinical effectiveness, health utility and cost and healthcare resource use were thoroughly reviewed. After sorting variables, they were aggregated if the information is minor and can be categorised into one variable. The areas where aggregation is mostly required are resource use in economic evaluation. In order to reflect the current practice, especially cost part has naturally incorporated RWD into

the analysis. It is less informative to collapse the variables in resource use since the use of RWD in cost part is already expected. Also, different health technology can include different resource use reflecting its characteristics. Counting every cost part is not an accurate way to understand why and how RWD was used. Hence, the variable, resource use is not differentiated. However, variables such as volume of treatment or dose adjustment are separated from resource use as these data have potentially critical impacts on the result of economic evaluation.

Step 3: Validation of data extraction tool

The data extraction tool will be validated by a second researcher independently repeating the data extraction for a random sample of appraisals (20% of all appraisals). This validation is required to check the replicability of the data extraction and the clarity of the extraction tool. Any disagreements between the researchers will be resolved by discussion. Peer discussion following the validation process is important not only to check the clarity of this protocol but also to investigate any deviations caused by unclear information. It will help pinpoint where a higher degree of subjectivity may arise in the data extraction.

Step 4: Data analysis

The extracted data will be analysed quantitatively in two different ways. First, a descriptive analysis will summarise where and how RWD has been used in appraisals. This will be supplemented by an analysis of the intensity of use of RWD in order to explore changes in the pattern of use of RWD over time and differences with respect to cancer type. In addition to descriptive statistics, the association between years and the intensity of use of RWD will be examined. Secondly, a regression analysis will be performed to investigate which factors are

associated with the greater use of RWD in a company's submission. A part of the protocol development, appraisal documents were reviewed to identify factors potentially associated with the use of RWD. Five factors were identified and formulated into hypotheses about increased use of RWD (Figure 3).

Methodological issues

The design of this data extraction protocol, in which information is reliably and repeatedly extracted across appraisals, will allow us to review evidence for the use of RWD more systematically than could be obtained from a number of case studies. However, several methodological challenges can be anticipated. This section addresses these challenges and how they might be mitigated.

Issue 1: Unclearly stated information

Overall, NICE appraisals clearly describe the data used in the evidence synthesis. However, in few cases, a result of systematic research is listed without clarification that the study is included in evidence synthesis. In appraisals, systematic literature review is carried out to identify all relevant evidence. Clinical effectiveness evidence is carefully examined and described in detail, with clear reasons for the inclusion and exclusion of studies. On the other hand, the systematic search for resource use and cost information usually enumerates miscellaneous studies with bibliographic information and a summary, but the critical review of minor components of health cost is sometimes missing. While a manufacturer provides the result of the assessment, it is possible that some manufacturers' submissions do not clearly state whether a particular study is used to determine which elements of resource use made

up the health state costs. However, it appears to be rare for there not be an explicit statement regarding the evidence used, mostly with respect to resource use.

Issue 2: Level of aggregation

An important question is the most appropriate level of aggregation. This is best illustrated with respect to healthcare costs. It would be possible to have a variable indicating use or non-use of RWD for every single element of cost (distinguishing GP visits, frequency of hospitalization, and so on). At the opposite extreme there could be a single cost variable which indicated whether RWD was used for any element of cost. The more aggregated the measure the greater the loss of information, but some elements of cost are much more important than others and the potential analyses of the use of RWD will multiply greatly if there is no attempt at aggregation. The current protocol tries to balance the advantages and disadvantages of different levels of aggregation by combining several elements into a health state cost variable but distinguishing other important components of cost such as, such as volume of treatment, dose adjustment and resource use for adverse events.

Issue 3: No consensus on the definition of RWD.

This research uses the definition of RWD merging definitions by FDA and Makady et al. The distinctive part of the definition used in this research is 'routinely collected' data from 'non-experimental study'. Although the definition provides specific and clear definition for this research, there is no consensus as how best to define RWD. Even same definition can be interpreted in a different way. For example, some researchers interpret that 'routinely collected' in FDA definition is 'collected in routine care' whereas other interpret it as 'how frequently data are collected.' It is likely that other definitions of RWD are preferred by other

researchers and the data extracted will be influenced by the definition of RWD chosen. It was considered to use multiple definitions of RWD. However, this research does not work with several definitions due to practical problems such as that they multiply the number of potential analyses and make data extraction take longer. Although the definition can be questioned by other researchers who have different views, the various definitions overlap considerably. It is thus unlikely there will be a marked divergence in the data extracted when using the different definitions.

Design to mitigate methodological issues

Several operational rules are designed to minimise bias likely to come from the methodological issues encountered in the data extraction. First, 'Unclear' is recorded separately in order to provide a more accurate description of the use of RWD. However, for purposes of data analysis, we anticipate treating these instance as "no RWD" since the code 'not clear' cannot be independently analysed. In addition, having a 'not clear' category in analysis is unlikely to improve data quality since this problem appears to arise in very few appraisals. Also, the information which is not clearly recorded in the appraisal documents is usually not major information with respect to the evidence synthesis. It is also closely linked to the reason for using binary code for analysis in this research. Decomposing levels of codes into several small parts can help to extract the information as it is. However, it is more likely to increase the complexity since trivial information is individually recorded. The extracted trivial data should be interpreted based on another operational rule. It is subject to increased error, particularly when testing hypotheses. For these reasons, the benefit of breaking up the level of codes into multilevel does not outweigh the benefit of binary codes while separation is much more time consuming. Instead of adapting multilevel codes, this study suggests an

alternative, an intensity analysis which helps to understand valid difference within diverse patterns of the use of RWD. When looking at the pattern of the use of RWD, the intensity of the use of RWD will be analysed. Simply counting the number of times RWD is used is not an accurate way to understand why and how RWD was used. Alternatively, this study focuses on variables which are more critically reviewed in appraisal. Variables such as survival outcome, volume of treatment and choice of comparators are more likely to influence estimated cost-effectiveness. Especially, the survival outcome is the most important information in both clinical and cost-effectiveness as well as one of the controversial areas where to use RWD. The intensity analysis is a framework to show whether RWD is used in these components alongside the quantity of the use of RWD. It can offer more benefits in deeper understanding of the use of RWD than counting all miscellaneous uses of RWD.

STRENGTHS AND LIMITATIONS

To the best of the authors' knowledge, this is the first study protocol to investigate to what extent RWD has been used in NICE appraisals. It allows the practice of extracting information to be reproducible, systematic and transparent. Strengthening the reproducibility and transparency of extracting process can maximise the understanding of the use of RWD by allowing more accurate interpretation and use of their findings. This protocol could be relevant to researchers or HTA agencies who aim to understand how various data resource was used in HTA in a context of England. The findings of this protocol can provide the full picture of the use of RWD in NICE appraisal over ten years. Moreover, the study findings could add more value under the discussion of NICE which tries to broaden the evidence in NICE practice.

The protocol has the limitation that it has been developed to study the use of RWD in NICE appraisals of oncology drugs. Consequently, the data extraction protocol may not be fully applicable to appraisals in other disease areas or to the different practice of other HTA bodies. Since the documentation is significantly different depending on each country's context, it is not feasible for the protocol to extract the same information as in English context. However, many of the distinctions are of wider application, e.g. parametric vs non-parametric use of RWD, and the taxonomy of where in an economic evaluation it might be relevant to look for use of RWD. Also, the hypotheses are potentially of wider application. The results are going to be specific to NICE but otherwise the structure of this research has wider application. Although not fully transferrable, this protocol can be modified for use in other HTA contexts.

Ethics and dissemination

This study is approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine on 14 November 2019 (17315). Results will be published in peer-reviewed journals.

PATIENT AND PUBLIC INVOLVEMENT

No patient involved.

ACKNOWLEGEDMENTS

The authors thank to David Lugo-Palacios for useful advice on clarity of data extraction and validation process.

CONTRIBUTORS

Both authors contributed to conceptualising and designing the study. JK drafted the protocol manuscript. JC revised the manuscript for important intellectual content and contributed to the methodology.

FUNDING STATEMENT

JK is supported by the Centre for Cancer Biomarkers, University of Bergen funded by the Research Council of Norway grant number (223250). The funder was not involved in the design of the protocol or decision to submit the protocol for publication, nor will they be involved in any aspect of the study conduct.

COMEPTING INTEREST STATEMENT

None declared.

FULL REFERENCES

1. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. Value Heal [Internet]. 2017;20(7):858–65. Available from: http://dx.doi.org/10.1016/j.jval.2017.03.008

2. WHO. WHO | HTA Definitions. WHO [Internet]. 2015 [cited 2019 May 13]; Available from: https://www.who.int/health-technology-assessment/about/Defining/en/

3. York Health Economics Consortium. Health Technology Assessment (online) [Internet]. 2016 [cited 2021 Oct 4]. Available from: https://yhec.co.uk/glossary/health-technology-assessment/

4. Chan K, Nam S, Evans B, Deoliveira C, Chambers A, Gavura S, et al. Developing a framework to incorporate real-world evidence in cancer drug funding decisions: The Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration. BMJ Open. 2020 Jan 7;10(1).

5. Hilgers R. Design and analysis of clinical trials for small rare disease populations. J Rare Dis Res Treat. 2016 Nov 1;1(3):53–60.

6. Averitt AJ, Weng C, Ryan P, Perotte A. Translating evidence into practice: eligibility criteria fail to eliminate clinically significant differences between real-world and study populations. npj Digit Med 2020 31 [Internet]. 2020 May 11 [cited 2021 Oct 5];3(1):1–10. Available from: https://www.nature.com/articles/s41746-020-0277-8

7. Latimer N. Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data - PubMed - NCBI. Med Decis Mak [Internet]. 2013 [cited 2019 May 14];33(6):743–54. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27905716

8. Davies C, Briggs A, Lorgelly P, Garellick G, Malchau H. The “hazards” of extrapolating survival curves. Med Decis Mak [Internet]. 2013 Apr [cited 2020 Aug 17];33(3):369–80. Available from: https://pubmed.ncbi.nlm.nih.gov/23457025/

9. Katkade VB, Sanders KN, Zou KH. Real world data: An opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. J Multidiscip Healthc. 2018;11:295–304.

10. Littlejohns P, Chalkidou K, Culyer AJ, Weale A, Rid A, Kieslich K, et al. National Institute for Health and Care Excellence, social values and healthcare priority setting [Internet]. Vol. 112, Journal of the Royal Society of Medicine. SAGE Publications Ltd; 2019 [cited 2021 May 11]. p. 173–9. Available from: https://journals.sagepub.com/doi/full/10.1177/0141076819842846

11. Schlander M. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Springer New York; 2007.

12. Broader types of data to be used in development of NICE guidance | News and features | News | NICE [Internet]. [cited 2020 May 5]. Available from: https://www.nice.org.uk/news/article/broader-types-of-data-to-be-used-in-development-of-nice-guidance

13. Wang S, Goring SM, Lozano-Ortega G. Inclusion of real-world evidence in submission packages to health technology assessment bodies: What do current guidelines indicate? Value Heal [Internet]. 2016 May 1 [cited 2020 Jun 22];19(3):A287. Available from: http://www.valueinhealthjournal.com/article/S1098301516008317/fulltext

14. Pietri G, Masoura P. Market Access and Reimbursement: The Increasing Role of Real-World Evidence. Value Heal [Internet]. 2014 Nov 1 [cited 2020 Jun 22];17(7):A450–1. Available from:

http://www.valueinhealthjournal.com/article/S1098301514031465/fulltext

15. Harwood M, Deighton A, Mickle A, Qian C, Szabo S. The use of real-world data in health technology assessment of medications for rare diseases. Value Heal [Internet]. 2019 Nov 1 [cited 2020 Sep 10];22:S863. Available from: http://www.valueinhealthjournal.com/article/S1098301519348296/fulltext

16. Roberts MH, Ferguson GT. Real-World Evidence: Bridging Gaps in Evidence to Guide Payer Decisions. PharmacoEconomics - Open [Internet]. 2021 Mar 18 [cited 2021 Jun 29];5(1):3–11. Available from: https://doi.org/10.1007/s41669-020-00221-y

17. Bullement A, Podkonjak T, Robinson MJ, Benson E, Selby R, Hatswell AJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. Int J Technol Assess Health Care [Internet]. 2020 Aug 1 [cited 2021 May 12];36(4):388–94. Available from: https://doi.org/10.1017/S0266462320000434

18. FDA. Real-World Evidence [Internet]. 2020 [cited 2020 May 6]. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

19. Guide to the multiple technology appraisal process [Internet]. [cited 2020 Aug 7]. Available from: www.nice.org.uk

Figure Legends

Figure 1 Inclusion/exclusion criteria

Figure 2 The framework for data extraction

Figure 3 Hypotheses about increased use of RWD

Supplement Legends

Supplement 1 Glossary of variables in extraction template

Inclusion criteria
- STA of oncology medicine
- Appraisals issued from January 2011 to May 2021
Exclusion criteria
- Appraisal of technology for preventing the complications of cancer
- Appraisal of surgical practice and other therapeutic therapies
- Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals)

Figure 1 Inclusion/exclusion criteria

131x35mm (240 x 240 DPI)

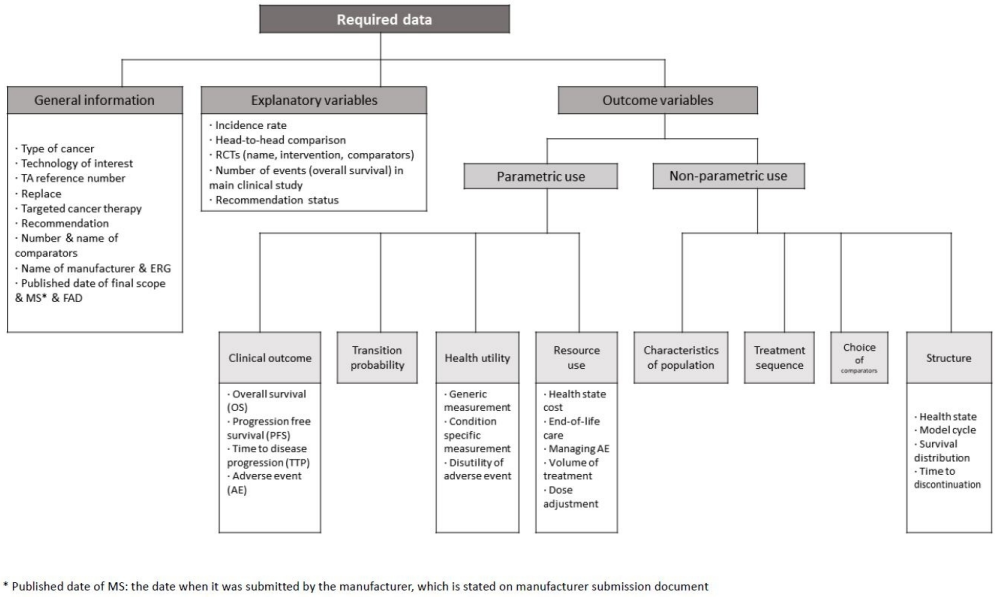


Figure 2 The framework for data extraction

141x84mm (240 x 240 DPI)

1) Poor internal/external validity of the clinical trial is associated with greater use of RWD.
2) Absence of direct (head-to-head) comparison is associated with greater use of RWD.
3) Low incidence rate of the disease is associated with greater use of RWD.
4) Immature survival data in the clinical trial are associated with greater use of RWD.
5) The technology having been recommended in previous NICE TA guidance is associated with greater use of RWD.

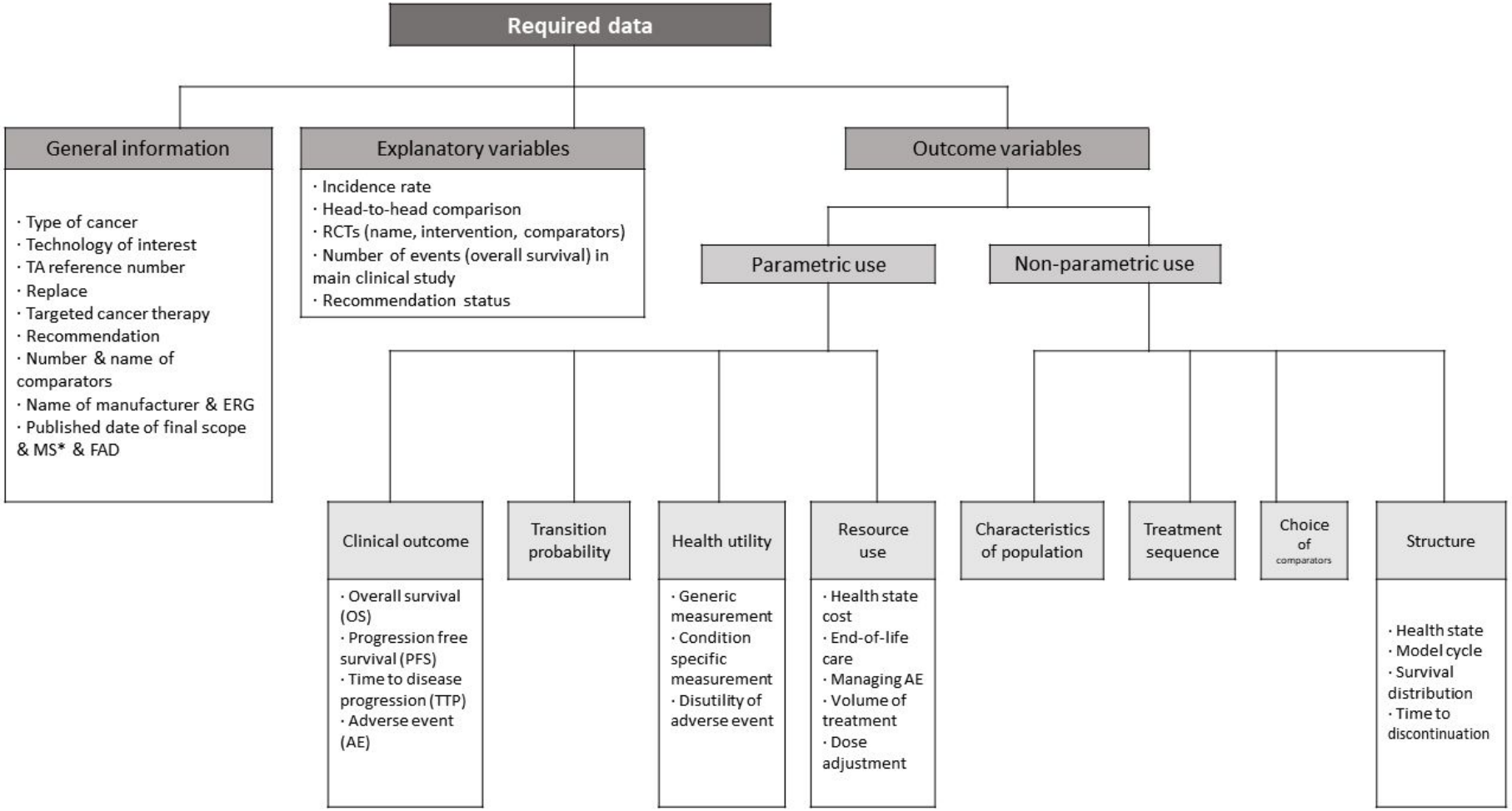
Figure 3 Hypotheses about increased use of RWD

93x24mm (240 x 240 DPI)

Figure 1 Inclusion/exclusion criteria

Inclusion criteria
- STA of oncology medicine
- Appraisals issued from January 2011 to May 2021
Exclusion criteria
- Appraisal of technology for preventing the complications of cancer
- Appraisal of surgical practice and other therapeutic therapies
- Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals)

Figure 2 The framework for data extraction



* Published date of MS: the date when it was submitted by the manufacturer, which is stated on manufacturer submission document

Figure 3 Hypotheses about increased use of RWD

1) Poor internal/external validity of the clinical trial is associated with greater use of RWD.
2) Absence of direct (head-to-head) comparison is associated with greater use of RWD.
3) Low incidence rate of the disease is associated with greater use of RWD.
4) Immature survival data in the clinical trial are associated with greater use of RWD.
5) The technology having been recommended in previous NICE TA guidance is associated with greater use of RWD.

Supplement 1 Glossary of variables in extraction template

General information		
Variable	**Explanation**	**Coding**
Type of cancer	The NICE classification of the cancer (website: https://www.nice.org.uk/guidance/conditions-and-diseases/cancer)	Bladder cancer=1, Blood and bone marrow cancer =2, Breast cancer=3, Colorectal=4, Neuroblastoma=5, Head and neck=6, Liver=7, Lung=8, Oesophageal=9, Ovarian=10, Pancreatic=11, Prostate=12, Renal=13, Skin=14, Stomach=15, Sarcoma=16
Technology of interest	The name of drug in the current appraisal. If it is combination therapy, the key technology which manufacturer focuses on will be taken here.	Narrative description
Indication	Clinical indications which are addressed in Final Appraisal Determination (FAD) document	Narrative description
TA number	the reference number of the technology guidance	Narrative description
Replace	Whether TA guidance has replaced or not. Appraisals can be replaced after rapid reviews/reviews/updates of previous appraisals or CDF reviews. Regardless of reasons of replacement, TA reference number which is replaced by this appraisal of interest will be recorded.	None= 0 If current appraisal replaces previous appraisal, the replaced TA reference number is recorded here.
• Pre-2016 CDF reconsideration	Before April 2016, the drug which was not reviewed or not recommended for routine commissioning by NICE can be used using the previous model of CDF. When new CDF was introduced in April 2016, these drugs in the old CDF were appraised by NICE to transit the model of CDF. This variable describe whether the appraisal of interest is an appraisal of the CDF reconsideration for the drug used in the old model of CDF before 2016.	No, it is not pre-2016 CDF reconsideration =0 Yes, it is a appraisal of pre-2016 CDF reconsideration =1
• 2016 CDF review	In April 2016, a new model of CDF was introduced. In the new model, an additional recommendation, recommended for use within the CDF is available when NICE appraising cancer drugs. The drug available via the CDF has to collect the data for further review for the routine commissioning after a certain period. As this mandated data collection can impact on the use of RWD, this variable allows to distinguish the appraisals, which RWD is more likely to be used.	No, it is not 2016 CDF review =0 Yes, it is 2016 CF review=1
Targeted cancer therapy	Treatment that uses drugs or other substances to identify and attack specific types of cancer cells	Non-targeted therapy = 0, targeted therapy = 1, not sure = Narrative description

Recommendation	the classification of recommendations made by the NICE committee in FAD document - Not recommended: 0 - Recommended (in line with marketing authorisation): 1 - Recommended (in line with marketing authorisation) in CDF:2 - Optimised: 3 - Optimised in CDF: 4 - Recommended in research: 5	Not recommended=0, recommended=1, recommended (cdf)=2, optimised=3, optimised (cdf)=4, recommended in research=5
number of comparators	Count the number of comparators in each manufacturer submission or FAD document. The information in manufacturer submission and FAD is recorded in the separated rows (manufacturer row/committee row).	Number in the manufacturer's submission
name of comparators	Record the name of comparators in manufacturer submission or FAD document	Narrative description
name of manufacturer	the name of manufacturer in manufacturer submission	Narrative description
name of the ERG	the name of the ERG (evidence review group)/AG (assessment group) in ERG critiques or AG reports	Narrative description
published date of final scope	the date of final scope as MM/YYYY	Date (MM/YYYY)
published date of manufacturer	the date of manufacturer submission as MM/YYYY.	Date (MM/YYYY)
published date of FAD guidance	the date of FAD document as MM/YYYY	Date (MM/YYYY)
Explanatory variables		
Variable	**Explanation**	**Coding**
Incidence (rate, year)	The rate would be recorded as it is in the appraisal. Incidence rate could be found in the final scope document or in manufacturer submission document. If the figures are not identical in each document, the latest rate is recorded. Most appraisals present the annual estimate of the number of patients who are eligible for the treatment in the "Budget Impact" section of company submission. This number is mainly used for the incidence. If this information is not available in the appraisal, the number in previous appraisal for similar indication is used instead.	Number
H2H	Whether the head-to-head clinical trial of a technology of interest exists or not, which compares with agreed comparators. The information is most likely to be found in the section: Identification and selection of relevant studies in clinical effectiveness part.	no=0, yes=1, yes but some comparators missing =2

• ITC	ITC (indirect treatment comparison). The information could be found in the section: Indirect and mixed treatment comparisons in clinical effectiveness part.	no=0, yes=1
• RCT (technology of interest)	Main RCT used in the appraisal: the name of the H2H RCT, if it exists. Unless there is an H2H, RCT refers to the clinical trial of technology of interest in the ITC.	no=0, yes=1
- Name of RCT	The name of the aforementioned RCT	Narrative description
- Intervention in RCT	The name of the intervention used in the aforementioned RCT. This variable helps to identify the main technology in RCT when technology is appraised as combination therapy.	Narrative description
- Comparators in RCT	The comparator of the aforementioned RCT	Narrative description
- Size of RCT	The number of participants in the aforementioned RCT	Number
- Median duration of follow-up	The median duration of follow-up in the aforementioned RCT. If it is not reported, record as NR (not reported).	Unit: month Not reported = ..
• Anchored/unanchored	"Anchored" means that RCT of technology of interest exists, and the RCT has been linked to any other studies which evaluate the drug's effectiveness. "Unanchored" means that the clinical outcome study doesn't have any comparators which connect to other studies. For example, comparing a single-arm study with a single-arm study is "unanchored". Also, RCTs compared without common comparators in ITC is "unanchored".	Not anchored=0, Anchored =1
• MAIC/STC	Matching adjusted indirect comparison (MAIC), Simulated Treatment Comparison (STC). A methodology of making adjustment to increase the comparability of two distinct populations mostly among unanchored studies. But it could be used in anchored studies in case where the two populations in ITC is starkly different from each other.	Naive=0, MAIC=1 STC=2 Other methods=3
Risk of bias (RoB) of RCT (direct quotation)	In order to evaluate the internal validity of RCTs, the risk of bias, which was reported in the ERG report, will be recorded here. Information is available at the quality assessment part of the ERG report. The ERG assesses the risk of bias of the included study using quality assessment tools. The ERG statement is directly quoted. The ERG often addresses the issue of quality of study narratively. Moreover, the ERG uses different terminology, whereas the domain of assessment is consistent. Therefore, the risk of bias would be narratively recorded. Prior to analysis, it will be scored by looking at the number of factors about which the ERG has expressed concern.	Direct quotation from ERG documents
• Risk of bias in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into four groups following the number of risk factors.	High/good quality without mentioned weakness= 0, risk factor 1 (low) =1, risk

		factor 2-3 (moderate)=2, risk factor 4 (high) =3
External validity of RCT	As narrative accounts, generalisability of RCT is reported in the ERG report whether the population of RCT properly represents the UK general population in terms of aging structure, health status and health care practice (practice-dose, subsequent treatment, etc.).	Direct quotation from ERG documents
• External validity in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into three groups following the severity of generalisability assess by ERG.	Representative without mentioned weakness= 0, Representative but minor concerns =1, Questionable generalisability =2
Previously recommended in other indication	Whether the technology has been recommended for other types of cancers besides the current indication of the technology.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in other indication	If it was recommended for other indications, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Previous recommended treatment in the same cancer	Whether the technology has been recommended for other treatment lines in the same type of cancer.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in the same cancer	If it was recommended for other treatment lines in the same cancer category, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Maturity of survival data in clinical trial	The data maturity is examined by looking at the number of events (deaths) of intervention arm in clinical trials. In published appraisal document, some of the information is redacted due to confidentiality. If the information is not available, the article of clinical trial published in journals is searched in order to check how many events are observed during the trial. Nonetheless, data are still not available in some cases. Since manufacturer is likely to redact the OS information when median OS was not reached. Hence, the survival data in this case are regarded as immature.	Direct quote from manufacturer submission
• Maturity (grade)	The direct quotation will be classified into three groups following the data cut point, 20% and 50 % of the number of events. This protocol adapts the criterion for measuring maturity of survival data in Tai et al. which investigates data maturity in STAs by looking at the proportion of death in pivotal trials. In the study, 20, 50 and 70 % of proportion of number of deaths are used to discuss the maturity of survival data (1). This protocol only uses 20% and 50% to assess the maturity without the category “unclear.”	Immature (number of events < 20%) =1, Relatively immature (20% ≤ number of events≤50%)=2 Mature (number of events < 50%) =3

Outcome variables			
Variable	**Explanation**	**Coding**	**Example**

characteristic of population	Whether RWD are used to determine the characteristic of population, including the initiation age and health performance status (ECOG) or not. - Soft use: when RWD are supplementary evidence to decide the population characteristics - Hard use: when RWD determine the characteristics of population in economic evaluation	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Pomalidomide, in combination with low - dose dexamethasone, for treating multiple myeloma in adults at third or subsequent relapse (NICE TA427): baseline patient characteristics were obtained from RWD collected from a hospital population since the majority of the trial populations were previously untreated, which was different from target population.
treatment sequence	Whether RWD are used to determine the subsequent treatment option or not. After the disease progression onto the later stages of cancer treatments, patients are likely to receive idiosyncratic subsequent treatments. The pattern of subsequent treatment for cost-effectiveness analysis could be observed by RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): a study of medical records was used to determine the treatment sequence.
choice of comparator	Whether RWD are used to choose the comparators in economic evaluation or not. Although comparators are chosen based on the current clinical guideline, drug utilisation data or clinical expert opinion are frequently referred to find the most relevant comparators in evaluation.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Ixazomib with lenalidomide and dexamethasone for treating relapsed or refractory multiple myeloma (NICE TA505): the manufacturer considered that lenalidomide was appropriate comparator based on IMS market research data (lenalidomide, 69% market share and panobinostat, 7%).
structure (health state)	Whether RWD are used to determine the health state such as stable, progression, and death in a given model. Information is available at health state in the model of cost-effectiveness analysis in manufacturer submission documents.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): the model health state of post-progression was specified based on a retrospective patient medical record review study.
structure (model cycle)	Whether RWD are used to determine model cycle or not. Model cycle, hereby, means that the duration between different health states, which can be influenced by the severity of conditions.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A **
Structure (survival distribution of intervention)	Whether RWD are used to decide the survival distribution of intervention or not. Since survival rate observed in RCTs is immature, it is necessary to extrapolate the survival rate for analysis. In order to choose proper survival distribution, the goodness of fit is tested (AIC, BIC).	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Larotrectinib for treating advanced solid tumours with NTRK fusions (NICE TA630): UK all-cause mortality data were used to assess the clinical acceptability of distributions whether patient overall survival exceeded current UK life expectancy

	Also, the clinical plausibility is asked to validate the distribution. In this case, the alternative data can be utilized. - If RWD is utilised for choosing distribution, mark as "hard use". - If RWD is utilised as supplementary evidence for the chosen distribution, mark as "soft use".		
Structure (survival distribution of comparator)	Whether RWD are used to validate the feasibility of survival distribution of comparator or not. As survival distributions of intervention and comparators are separately determined, the extraction tool approach it independently. Apply the abovementioned description on survival distribution of intervention to comparator in this row.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	
Structure (Time to discontinuation of intervention)	Whether RWD are used to decide the time to discontinuation of intervention or not. The time to discontinuation is likely to be decided by 1) simply adopting discontinuation rule in trials, 2) formulating distribution of discontinuation, or 3) clinical experts' opinion. - If RWD are used for designating the time to discontinuation, mark as "hard use" - If RWD are used as supplementary evidence for designating the time to discontinuation, mark as "soft use". - If clinical experts' opinions are used for designating the time to discontinuation, it is not regarded as RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Lorlatinib for previously treated ALK-positive advanced non-small-cell lung cancer (NICE TA628): The plausibility of the extrapolation of time on treatment was validated by UK RWD, hospital network data.
Structure (time to discontinuation of comparator)	Whether RWD are used to decide the time to discontinuation of comparator or not. Apply the above-mentioned description on time to discontinuation of intervention to comparator in this	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	

	row.		
Clinical outcome (OS) intervention	Whether RWD give the figure for overall survival (OS) of intervention or not. In order to measure the Quality Adjusted Life-Years (QALYs), it is necessary to extrapolate overall survival based on observed data on survival. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Nivolumab for adjuvant treatment of completely resected melanoma with lymph node involvement or metastatic disease (NICE TA558): the survival model applied the registry data (American Joint Committee on Cancer; AJCC) to both treatment arms after a certain time point.
Clinical outcome (PFS) intervention	Whether RWD give the figure for progression free survival (PFS) of intervention or not. The progression of disease is important for economic evaluation model in terms of health state transitions and treatment switching. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (RR) intervention	Whether RWD provides the response rate (RR) for the intervention or not. The effectiveness of cancer treatment is often shown by responses of tumour cells, which is evaluated by the RECIST criteria or other criteria. The response rate data would be collected in RCT or other type of data.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (TTP) intervention	Whether RWD give the figure for time-to-progression (TTP) of intervention or not. Some cancer treatments show their clinical effectiveness not through the progression free survival (PFS), but alternatively through time-to-progression.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (AE) intervention	Whether RWD give the figure of adverse event (AE) of intervention or not. Adverse events are crucial information for the estimation of the QALYs. The adverse events are collected in RCT. However, RWD, including cohort studies, retrospective studies, or other type of studies, also provide the information of adverse events, which cannot be found in RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Blinatumomab for treating acute lymphoblastic leukaemia in remission with minimal residual disease activity (NICE TA589): retrospective non-interventional cohort study collected from 2000 to 2017 was used to inform the clinical outcome of comparators as well as adverse event.
Clinical outcome (OS) comparators	Whether RWD give the figure of overall survival (OS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, clinical outcome (OS) intervention

Clinical outcome (PFS) comparators	Whether RWD give the figure for the progression free survival (PFS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (RR) comparators	Whether RWD provide the response rate (RR) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (TTP) comparators	Whether RWD provide the time-to-progression (TTP) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (AE) comparators	Whether RWD provide the figure adverse events (AE) for the comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, clinical outcome (AE) intervention
Transition probability	Whether RWD provide the transition probability from one state to other state, if it is applicable.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Pembrolizumab for treating melanoma with high risk of recurrence (NICE TA553): electronic health records (Flatiron database) collected by cancer care providers in the US was used to model transition from the "locoregional recurrence (LR)" state to the "distant metastases" and life tables for transition from the LR to "death" state.
Health utility of health state (generic)	Whether health state utility survey of generic measurement is done in RWD or RCT. Health state utility is necessary information for the estimation of the QALYs. Generic health utility measurement, EQ-5D, is frequently used. There is national tariff of EQ-5D to get the scores. Hereby, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Health utility of health state (condition-specific)	Whether health state utility survey of condition-specific measurement is done in RWD or RCT. In cancer treatment, condition-specific measurement is commonly adopted. Similar to the previous row, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**

Disutility of adverse events	Whether survey of collecting disutility data is done in RWD or RCT. As adverse events are likely to reduce the patient's quality of life, the disutility of adverse events is included in estimates. The way of collecting survey (RWD or RCT) is drawn to attention.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Health state cost) common	Whether resource use for estimating health state cost is derived from RWD or RCT. In economic evaluation, the unit cost mostly comes from the national reference cost. The total cost is calculated by the total resource use (volume of technology and health care services) multiplied by the reference cost. Here, the only resource use is focused in data extraction. Resource use for estimating health state cost includes all activity like monitoring, GP visits, pharmacy cost etc. Health state resource use could be aggregated or individually listed. Here, the difference of describing health state cost is not separately considered.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies (NICE TA559): RWD was used for estimating the cost of inpatient admission (data: Hospital Episode Statistics), the cost of home care and hospice (data: National Audit Office), and GP time (data: Personal Social Services Research Unit; PSSRU).
Resource use (end-of-life care) common	Whether resource use for estimating end-of-life care is derived from RWD or RCT. Resource use of terminal cancer patients is not frequently reported in the RCT providing the treatment effect. Therefore, other data resources, including RCTs of other technologies, provide the information of resource use in the end-of-life care.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Managing AE) intervention	Whether resource use for managing adverse events of intervention is derived from RWD or RCT. Resource use of managing adverse events is reported in RCTs as well as in other types of researches which can provide alternative perspectives.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (volume of treatment) intervention	Whether resource use for volume of treatment of intervention is derived from RWD or RCT. In this study, scope of the volume of treatment is limited to the frequency of treatment, frequency of administration, and amount of subsequent	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Fulvestrant for treating untreated locally advanced or metastatic oestrogen-receptor positive breast cancer (NICE TA503): a medical chart review study was used to determine the proportion of patient using subsequent treatment for cost calculation.

	treatment.		
Resource use (Dose adjustment) intervention	Whether resource use for dose adjustment of intervention is derived from RWD or RCT. There are several reasons for adjusting dose such as adverse events (AEs). The dose of cancer treatments is calculated by BSA (body surface area). This study focuses only on BSA and dose adjustment due to AEs, because these information are commonly reported in NICE appraisals.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Managing AE) comparators	Whether resource use for managing adverse events of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (volume of treatment) comparators	Whether resource use for volume of treatment of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, resource use (volume of treatment) intervention
Resource use (Dose adjustment) comparators	Whether resource use for dose adjustment of comparators is derived from RWD or RCT. Since the intervention is a novel technology, RCTs provide less information on the adjustment. RWD could be utilised to provide more relevant information regarding dose adjustment of existing technologies which have been used in routine clinical practice.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**

* In order to detect the use of RWD in sensitivity analysis, the parametric part is duplicated.
** As data extraction is not conducted, all of examples are not available at this stage. In this case, it marked as N/A.
*** Benefits/challenges of the use of RWD are collected in outcome variables.
**** In cases where trials have more than two arms, only the arms considered as relevant for decision problem in evidence submission are included. If there are two intervention arms and these arms are separately used for different indications in appraisals, the data extraction is carried out separately. When two arms are relevant as comparators for same indication, the data are recorded without distinguishing these arms.

1. Tai TA, Latimer NR, Benedict A, Kiss Z, Nikolaou A. Prevalence of Immature Survival Data for Anti-Cancer Drugs Presented to the National Institute

for Health and Care Excellence and Impact on Decision Making. Value Heal. 2020 Dec 8;

BMJ Open

A protocol for data extraction: how real-world data have been used in the National Institute for Health and Care Excellence appraisals of cancer therapy

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055985.R2
Article Type:	Protocol
Date Submitted by the Author:	13-Dec-2021
Complete List of Authors:	Kang, Jiyeon; London School of Hygiene and Tropical Medicine Faculty of Public Health and Policy, Department of Health Services Research and Policy; Cairns, John; London School of Hygiene and Tropical Medicine, Department of Health Services Research and Policy
Primary Subject Heading:	Health economics
Secondary Subject Heading:	Health economics
Keywords:	ONCOLOGY, HEALTH ECONOMICS, HEALTH SERVICES ADMINISTRATION & MANAGEMENT

A protocol for data extraction: how real-world data have been used in the National Institute for Health and Care Excellence appraisals of cancer therapy

Jiyeon Kang[1], John Cairns[1]

[1]Department of Health Service Research and Policy,

Faculty of Public Health and Policy,

London School of Hygiene and Tropical Medicine, London, UK

Corresponding author: Jiyeon Kang (Jiyeon.Kang@lshtm.ac.uk)

15-17 Tavistock place, London, WC1H 9SH, UK

Word count: 4,705

Keywords: Real-world data (RWD), Health Technology assessment (HTA), economic evaluation, oncology medicine, National Institute for Health and Care Excellence (NICE)

ABSTRACT

Introduction

Due to the limitations of relying on randomised controlled trials, the potential benefits of real-world data (RWD) in enriching evidence for health technology assessment (HTA) are highlighted. Despite increased interest in RWD, there is limited systematic research investigating how RWD has been used in HTA. The main purpose of this protocol is to extract relevant data from National Institute for Health and Care Excellence (NICE) appraisals in a transparent and reproducible manner in order to determine how NICE has incorporated a broader range of evidence in the appraisal of oncology medicines.

Methods and analysis

The appraisals issued between January 2011 to May 2021 are included following inclusion criteria. The data extraction tool newly developed for this research includes the critical components of economic evaluation. The information is extracted from identified appraisals in accordance with extraction rules. The data extraction tool will be validated by a second researcher independently. The extracted data will be analysed quantitatively to investigate to what extent RWD has been used in appraisals. This is the first protocol to enable data to be extracted comprehensively and systematically in order to review the use of RWD.

Ethics and dissemination

This study is approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine on 14 November 2019 (17315). Results will be published in peer-reviewed journals.

STRENGTHS AND LIMITATIONS OF THIS STUDY

Strengths

- This protocol enables data to be extracted in a transparent and systematic manner for the study of how RWD has been used in NICE appraisals including all the different ways an economic evaluation might use RWD.

- This study facilitates systematic understanding of the use of RWD in NICE appraisals over the last 10 years.

Limitations

- Since it is focussed on cancer, the methods and eventually the findings are to some extent

cancer-specific.

- The protocol could be modified to reflect the HTA context in different countries although the extraction protocol is not fully applicable to the practice of other HTA bodies as much of the protocol reflects the NICE appraisal process.

- Since data extraction is based on the four main types of appraisal document it is possible, but not likely that some relevant information concerning RWD is missed.

INTRODUCTION

In the last few years, interest in real-world data (RWD) has grown in health care decision-making (1). Health Technology Assessment (HTA) refers to the systematic evaluation of clinical- and cost-effectiveness of health technology (2,3). Health technologies include drugs, medical devices, diagnostics, surgical procedures to mitigate health issues and improve the quality of life (4). HTA requires valid and reliable information to evaluate such technoglogies. Randomised controlled trials (RCTs) have mainly provided the information (5). However, it is challenging to meet all information needs from RCTs since the new generation of therapies pose several assessment challenges. For example, when treatment options are expanding rapidly, it is increasingly unlikely that there are RCTs featuring of all the relevant comparators. Furthermore, the traditional design of RCTs is possibly less appropriate for new technologies such as those targeting rare genetic mutations where it is harder to recruit patients from the clinically relevant populations (6). Moreover, RCTs often have strict inclusion criteria reducing generalisability (7). Another barrier to obtaining the information required for HTA from RCTs relates to the extrapolation of survival. Extrapolation is required in order to incorporate the

survival data from RCTs in the health economic model (8). It is more challenging to identify the most appropriate extrapolation the shorter the duration of the trial. If survival data from RCTs are based on a very limited observation period, the extrapolation of the survival curve is likely to fail to predict the long-term effect (9).

The potential benefits of RWD in enriching evidence for HTA are highlighted by the limitations of relying on RCTs (10). This research focuses on the use of RWD in HTA by the National Institute for Health and Care Excellence (NICE). NICE has achieved an international reputation for rigorous development and application of scientific methods to appraise new health technologies to provide its decisions with robust and fair justification (11). More importantly, NICE is noted for the transparency of its processes, responsiveness to change, and commitment to using the best available evidence (12). The structure of the relevant documents facilitates identification of the key information and the documents are available on the NICE website. Therefore, review of these appraisals can provide comprehensive information on the evidence used for decision-making. In April 2020, NICE signalled its intention to integrate broader types of data in developing NICE guidance (13). Although it is primarily a statement of intent, it is not a new development in NICE practice since NICE already incorporates a diverse range of published scientific evidence when developing its guidance on health technologies. For example, UK audit data (TA255, 2012), Hospital Episode Statistics (TA559, 2018) and registry data such as the Edinburgh Ovarian Cancer Database (TA598, 2019), Surveillance, Epidemiology, and End Result program (TA562, 2019) have been used in the development of NICE technology appraisal (TA) guidance. While a wide range of data are already used in NICE guidance, there is limited understanding regarding how and where RWD has been used, and in which circumstances RWD is accepted as relevant.

Research is required to investigate systematically patterns in the use of RWD and to understand the driving forces behind its use in NICE appraisals.

Several researchers have reviewed practice across HTA bodies (14,15) or reported the use of RWD in HTA (16). However, little systematic research has been conducted. Important information is missing such as how they included literatures without selection bias, which parts of the evidence were reviewed, whether they have clearly defined RWD and justified or explained why this definition is relevant and how different HTA systems were compared given their different practices. Roberts et al. addressed the potential role of RWD in bridging the evidence gaps (17). However, they illustrate the use of RWD with a few examples, rather than providing a fuller picture of current practice when using RWD. Bullement et al. recently reviewed how RWD informed single technology appraisals of cancer drugs in NICE (18). Although this study follows a more systematic approach to the review of the use of RWD, a data extraction table was not provided and the authors focused only on how RWE influenced the cost-effectiveness analysis, and not on how RWE was used to support or establish the appraisal. Due to limited information presented concerning the review process in this study, it is unclear whether the information presented provides a full picture of the use of RWD. Bullement et al. included 113 STAs issued between April 2011 and October 2018. As interest in RWD is increasing over time, it may miss relevant information from recent years. This extraction protocol is required to help extract the data systematically from appraisals, to increase the reliability of the results of the analysis and to permit a more detailed description of the use of RWD and analysis of factors influencing its use.

A protocol is required to ensure the consistency of data extraction so that the risk of unsystematic data collection is reduced. The main purpose of this protocol is to extract data

from NICE appraisals in a transparent and reproducible manner to answer, "how has NICE incorporated a broad range of evidence in the appraisal of oncology medicines." Without proper justification and operational rules, the data may not be extracted consistently, with a risk of biasing the analysis. The extracted data are expected to be objective and less biased. By consolidating these data, subsequent analysis can provide more robust answers to questions regarding how RWD has been used in NICE technology appraisals. Furthermore, this protocol facilitates the development of a rich dataset which can highlight not just where RWD has been used but also what types of evidence have been used in the HTA process in line with NICE's interest in incorporating a broad range of evidence. The data can be analysed to answer several research questions including "how has RWD been used in NICE appraisals" and "which factors are associated with increased likelihood of the use of RWD" in depth.

METHODS AND ANALYSIS

NICE appraisal documents are identified following inclusion criteria (figure 1). The information is extracted from identified appraisals in accordance with extraction rules. The detailed extraction rules can be found in supplement 1. The extraction tool includes evidence-related information such as characteristics of the main clinical evidence and the economic evaluation model and other information. Using this tool, information will be collected about which parts of the cost-effectiveness analyses used RWD. Analyses of the intensity of use of RWD and regression analyses are planned. The data analysis is expected to start from January 2022 and be completed by December 2022.

Definition of RWD

A definition of RWD is clearly required before extracting information about the use of RWD in NICE. RWD is an umbrella term which covers broad categories of data. Although RWD is increasingly addressed in the literature, there is no consensus over the definition. One of the commonly used definitions of RWD is that of the US Food and Drug Administration (FDA)(19). Another widely cited study regarding the definition of RWD is Makady et al. (1). Each definition has relatively large operational flexibility to be used for data extraction. For example, companies sometimes present phase 1 clinical trial as RWD. However, these data hardly provide insights in the discussion of the use of RWD in HTA. Requiring data to meet both definitions can help to reduce the discretionary interpretation of RWD. Hence, this study uses a definition combining a category of the study designs of collecting RWD explored by Makady and his colleagues' study and the FDA's definition of RWD focusing on routinely collected data. In this research, RWD is defined as the data relating to patient health status and/or the delivery of health care routinely collected from non-experimental settings.

Step 1 Appraisal selection

The first step of the research identifies the NICE TA guidance which meets the eligibility criteria. TA guidance are publicly available on the NICE website (www.nice.org.uk). This study focuses on four types of appraisal documents, the final scope, the manufacturer's submission, the evidence review group (ERG) report, and the final appraisal determination. These documents are reviewed to establish whether RWD is used to determine any components of the economic evaluation.

Data sources

This research exclusively includes single-technology appraisals (STA) of oncology medicines. Figure 1 shows the inclusion and exclusion criteria. One aim is to understand how and where RWD has been used in the appraisal process. Therefore, it is necessary that the appraisal process should be identical. However, the STA and multiple technology appraisal (MTA) processes differ substantially. The MTA has different format of appraisal documents to assess several drugs or treatments used for one or more condition. It is challenging to gather the same information in the MTA process as different actors are responsible for producing and reviewing the main pieces of evidence (20). Besides, STAs are the predominant form in practice, 93% of appraisals of oncology. The small number of the MTAs, only eighteen oncology appraisals, limits the scope for a comparison of MTAs and STAs in terms of the use of RWD. Therefore, this study focuses on STAs, which assess a single treatment. It also limits analysis to appraisals published between January 2011 and May 2021 in order to have a long enough time period to capture potential changes over time in how RWD has been used but also recognising that STAs from earlier years might be of less interest because enthusiasm for RWD was largely absent. Here, the date when guidance was published refers to the date of issuing the final appraisal determination document (FAD) which can be regarded as an end point of the evidence synthesis process (in the absence of a successful appeal).

Operational separation

Following the inclusion and exclusion criteria, appraisals are identified. Among these appraisals, some TAs have more than one clinical indication or involve combination therapy. It is possible that different evidence was used for the different patient populations in the

appraisal. Hence, these appraisals are separated by clinical conditions or treatment lines and reviewed in order to avoid losing information. For example, olaparib for maintenance treatment of relapsed platinum-sensitive ovarian, fallopian tube or peritoneal cancer (NICE TA620) has two separate recommendations for different indications. While a patient who has a BRCA1 or BRCA2 mutation and has had three or more courses of platinum-based chemotherapy is eligible for the treatment, a patient who has a BRCA1 or BRCA2 mutation and has had two courses of platinum-based chemotherapy is able to use the treatment within Cancer Drug Fund. Consequently, these indications are included separately in the analysis.

Step 2: Data extraction

A detailed protocol is developed to guide the extraction of essential data for each appraisal in order to investigate the use of RWD in NICE technology appraisals in a systematic and reproducible manner. The protocol is designed to extract information from both the manufacturer's submission (manufacturer's cost-effectiveness analysis) and the final appraisal document (the model preferred by the committee) regarding where RWD was used, and to determine the extent to which the committee supported the use of RWD in these appraisals and understand what factors are associated with supporting or not supporting their use. Figure 2 shows the structure of the data extraction template. In summary, the extraction tool consists of three parts – general information, explanatory variables, and outcome variables. The outcome of interest being the use of RWD. The outcome variables record use or non-use of RWD for different elements of the economic evaluation. The information in the base-case analysis and sensitivity analyses will separately extracted. The tool includes all important elements of an economic evaluation. The study will analyse the data to investigate

patterns in the use of RWD in NICE appraisals, and the association between several factors and the use of RWD. Explanatory variables are suggested based on the hypotheses presented under Step 4: data analysis. All items in the extraction template and how to code them are described in the glossary (supplement 1). To convey the type of information to be extracted, some examples from a preparatory review are presented in the glossary.

Parametric and non-parametric use

This protocol distinguishes two categories of outcome variable, parametric and non-parametric use of RWD. Parametric use of RWD is the use of such data to define the numerical value of a specific variable in the economic evaluation, whereas non-parametric use is where data are utilised to develop the model structure or to determine the scope of the evaluation. For example, when RWD are used to estimate survival, this will be counted as parametric use with respect to clinical outcomes (OS/PFS). Parametric use is reviewed and recorded for the intervention and comparators separately as different data could be used in the cost-effectiveness analysis. An example of non-parametric use of RWD can be found in the appraisal of palbociclib for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495). In this appraisal, the company used information from a study of medical records to determine the subsequent treatments to be assumed in the economic model. This case is regarded as non-parametric use since RWD was used to specify the treatment sequence but not the quantity and cost of subsequent treatment.

Parametric and non-parametric use of RWD and the different categories shown in figure 2,

facilitate more consistent data extraction by highlighting the different ways RWD might be used, and provide greater flexibility when testing hypotheses regarding the use of RWD, and the exploration of ways to measure the intensity of use of RWD.

Coding

A key issue with respect to improving the reliability of data extraction is how many distinct variables to identify and how finely to divide are the potential responses to these variables. One option, in order not to lose information, is to have many distinct variables with binary responses. Another option is to merge many variables but have multi-level responses. This coding system has advantages which include avoiding information loss, and also grouping together 'similar' information used during appraisals to establish patterns of the use of RWD. This is closely linked to the reason for not using multiple responses in the coding. The template takes an "including all and combining trivia" approach. It helps to include all relevant variables where RWD data can potentially be used, but also to list variables more concisely by merging unnecessarily trivial variables so that the outcome of the extraction can be concretely analysed. Based on two categories, the parametric and non-parametric use of RWD, the areas where data are likely to be used are carefully searched. As a backbone of the extraction structure, distinguishing two categories helped to search each component systematically. Under parametric use, clinical effectiveness, health utility and cost and healthcare resource use were thoroughly reviewed. After sorting variables, they were aggregated if the information is minor and can be categorised into one variable. The area where aggregation is mostly required is resource use. In order to reflect routine clinical practice, especially the cost part has naturally incorporated RWD into the analysis. Estimates of unit costs are usually informed by NHS reference costs (a form of RWD) and thus in order

to provide a more sensitive measure of the use of RWD the extraction template focuses on resource use (with respect to cost). However, the measures of resource use are not fully differentiated. Different health technologies include different elements of resource use reflecting their characteristics. Distinguishing all resource use is not an accurate way to understand why and how RWD was used. Although all individual resource uses are not identified, some resource uses, which can be critical in appraisals are differentiated. Variables such as volume of treatment or dose adjustment have potentially critical impacts on the result of economic evaluation. Therefore, these variables are separated from overall resource use.

Step 3: Validation of data extraction tool

The data extraction tool will be validated by a second researcher independently repeating the data extraction for a random sample of appraisals (20% of all appraisals). This validation is required to check the replicability of the data extraction and the clarity of the extraction tool. Any disagreements between the researchers will be resolved by discussion. Peer discussion following the validation process is important not only to check the clarity of this protocol but also to investigate any deviations caused by unclear information. It will help pinpoint where a higher degree of subjectivity may arise in the data extraction.

Step 4: Data analysis

The extracted data will be analysed quantitatively in two different ways. First, counts and proportions will summarise where and how RWD has been used in appraisals. This will be supplemented by an analysis of the intensity of use of RWD in order to explore changes in the pattern of use of RWD over time and differences with respect to cancer type. In addition to descriptive statistics, the association between years and the intensity of use of RWD will be

examined. Secondly, a regression analysis will be performed to investigate which factors are associated with the greater use of RWD in a company's submission. As part of the protocol development, some appraisal documents were reviewed to identify factors potentially associated with the use of RWD. Five factors were identified and formulated into hypotheses about increased use of RWD (Figure 3).

Methodological issues

The design of this data extraction protocol, in which information is reliably and repeatedly extracted across appraisals, will allow us to review evidence for the use of RWD more systematically than could be obtained from conducting several case studies. However, several methodological challenges can be anticipated. This section addresses these challenges and how they might be mitigated.

Issue 1: Unclearly stated information

Overall, NICE appraisals clearly describe the data used in the evidence synthesis. However, sometimes the search process may not be well-documented and the precise source of information may not be clear. Systematic literature reviews are carried out to identify all relevant evidence in appraisals. Clinical effectiveness evidence is carefully examined and described in detail, with clear reasons for the inclusion and exclusion of studies. On the other hand, the systematic search for resource use and cost information usually enumerates miscellaneous studies with bibliographic information and a summary, but the critical review of minor components of health cost is sometimes missing. While manufacturers provide the result of their assessments, some manufacturers' submissions do not clearly state whether a

particular study was used to determine an element of resource use making up the health state costs. However, it appears to be rare for there not be an explicit statement regarding the evidence used, mostly with respect to resource use.

Issue 2: Level of aggregation

An important question is the most appropriate level of aggregation. This is best illustrated with respect to healthcare costs. It would be possible to have a variable indicating use or non-use of RWD for every single element of cost (distinguishing GP visits, frequency of hospitalisation, and so on). At the opposite extreme there could be a single cost variable which indicated whether RWD was used for any element of cost. The more aggregated the measure the greater the loss of information, but some elements of cost are much more important than others and the potential analyses of the use of RWD will multiply greatly if there is no attempt at aggregation. The current protocol tries to balance the advantages and disadvantages of different levels of aggregation by combining several elements into a health state cost variable but distinguishing other important components of cost, such as volume of treatment, dose adjustment and resource use for adverse events.

Issue 3: No consensus on the definition of RWD.

This research uses a definition of RWD merging definitions from the FDA and Makady et al. The distinctive part of the definition used in this research is 'routinely collected' data from a 'non-experimental study'. Although this definition provides a specific and clear definition for this research, there is no consensus on the best definition of RWD. Even the same definition can be interpreted in different ways. For example, some researchers interpret that 'routinely collected' in the FDA definition is 'collected in routine care' whereas other interpret it as 'how

frequently data are collected.' It is likely that other definitions of RWD are preferred by other researchers and the data extracted will be influenced by the definition of RWD chosen. While the use of multiple definitions of RWD was considered, it would create practical problems such as multiplying the number of potential analyses and making data extraction take longer. Although the chosen definition can be questioned by other researchers who have different views, the various definitions overlap considerably. It is thus unlikely there will be a marked divergence in the data extracted when using the different definitions.

Design to mitigate methodological issues

Several operational rules have been designed to minimise bias likely to come from the methodological issues encountered in the data extraction. First, 'not clear' is recorded separately in order to provide a more accurate description of the use of RWD. However, for purposes of data analysis, we anticipate treating these instance as "no RWD" since the code 'not clear' cannot be independently analysed. In addition, having a 'not clear' category in analysis is unlikely to improve data quality since we anticipate that this problem will arise in very few appraisals. Also, information which is not clearly recorded in the appraisal documents is usually not important information with respect to the evidence synthesis. The approach (extracting all relevant information which can provide meaningful data for analyses) is also closely linked to the reason for using binary code for analysis in this research. Decomposing levels of codes into several small parts can facilitate data extraction. However, it is more likely to increase the complexity since trivial information is individually recorded. The extracted trivial data should be interpreted based on another operational rule. It is subject to increased error, particularly when testing hypotheses. For these reasons, the benefit of using multi-level codes does not outweigh the benefit of binary codes while

separation is much more time consuming. Instead of adapting multilevel codes, this study will adopt an alternative approach, an intensity analysis which helps to identify important differences within the diverse patterns of use of RWD. When looking at the pattern of use of RWD, the intensity of use will be analysed. Simply counting the number of times RWD are used is not an accurate way to understand why and how RWD were used. Alternatively, this study focuses on variables which are potentially important determinants of cost-effectiveness in appraisal. Variables such as survival outcome, volume of treatment and choice of comparators are more likely to influence estimated cost-effectiveness. Especially, the survival outcome is the most important information in both clinical and cost-effectiveness as well as one of the controversial areas where to use RWD. The intensity analysis is a framework to show whether RWD is used in these components alongside the quantity of the use of RWD. It can offer more benefits in deeper understanding of the use of RWD than counting all miscellaneous uses of RWD.

STRENGTHS AND LIMITATIONS

To the best of the authors' knowledge, this is the first study protocol to investigate to what extent RWD has been used in NICE appraisals. It allows the practice of extracting information to be reproducible, systematic and transparent. Strengthening the reproducibility and transparency of data extraction can maximise understanding of the use of RWD by allowing more accurate interpretation and use of findings. This protocol could be relevant to researchers or HTA agencies who aim to understand how various data resources are used in HTA in England. Analysis of data generated using this protocol can provide a detailed picture of the use of RWD in NICE appraisals over ten years. Moreover, the study findings could add

value to NICE's ongoing work to broaden the evidence used in appraisals.

The protocol has the limitation that it has been developed to study the use of RWD in NICE appraisals of oncology drugs. Consequently, the data extraction protocol may not be fully applicable to appraisals in other disease areas or to the different practice of other HTA bodies. Since the documentation is significantly different depending on each country's context, it may not be feasible to extract the same information as in the English context. However, many of the distinctions are of wider application, e.g. parametric vs non-parametric use of RWD, and the taxonomy of where in an economic evaluation it might be relevant to look for use of RWD. Also, the hypotheses are potentially of wider application. The results are going to be specific to NICE but otherwise the structure of this research has wider application. Although not fully transferrable, this protocol can be modified for use in other HTA contexts. Lastly, this protocol focuses on four main documents. Relevant RWD may arise at the clarification or technical engagement stage. It is possible there is some information regarding use of RWD that is not reported in any of the four main documents. However, only a small number of such cases are anticipated. If RWD is critically used in a revised model and the committee thinks it is an important change, this evidence is likely to be addressed in FAD.

Ethics and dissemination

This study was approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine on 14 November 2019 (17315). Results will be published in peer-reviewed journals.

PATIENT AND PUBLIC INVOLVEMENT

None.

ACKNOWLEGEDMENTS

The authors thank David Lugo-Palacios for useful advice on the clarity of data extraction and the validation process.

CONTRIBUTORS

Both authors contributed to conceptualising and designing the study. JK drafted the protocol manuscript. JC revised the manuscript for important intellectual content and contributed to the methodology.

FUNDING STATEMENT

JK is supported by the Centre for Cancer Biomarkers, University of Bergen funded by the Research Council of Norway grant number (223250). The funder was not involved in the design of the protocol or the decision to submit the protocol for publication, nor will they be involved in any aspect of the study conduct.

COMEPETING INTEREST STATEMENT

None declared.

FULL REFERENCES

1. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. Value Heal [Internet]. 2017;20(7):858–65. Available from: http://dx.doi.org/10.1016/j.jval.2017.03.008

2. WHO. WHO | HTA Definitions. WHO [Internet]. 2015 [cited 2019 May 13]; Available from: https://www.who.int/health-technology-assessment/about/Defining/en/

3. York Health Economics Consortium. Health Technology Assessment (online) [Internet]. 2016 [cited 2021 Oct 4]. Available from: https://yhec.co.uk/glossary/health-technology-assessment/

4. Health technologies. 2021 Dec 10 [cited 2021 Dec 10]; Available from: https://www.euro.who.int/en/health-topics/Health-systems/health-technologies-and-medicines/health-technologies

5. Chan K, Nam S, Evans B, Deoliveira C, Chambers A, Gavura S, et al. Developing a framework to incorporate real-world evidence in cancer drug funding decisions: The Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration. BMJ Open. 2020 Jan 7;10(1).

6. Hilgers R. Design and analysis of clinical trials for small rare disease populations. J Rare Dis Res Treat. 2016 Nov 1;1(3):53–60.

7. Averitt AJ, Weng C, Ryan P, Perotte A. Translating evidence into practice: eligibility

criteria fail to eliminate clinically significant differences between real-world and study populations. npj Digit Med 2020 31 [Internet]. 2020 May 11 [cited 2021 Oct 5];3(1):1–10. Available from: https://www.nature.com/articles/s41746-020-0277-8

8. Latimer N. Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data - PubMed - NCBI. Med Decis Mak [Internet]. 2013 [cited 2019 May 14];33(6):743–54. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27905716

9. Davies C, Briggs A, Lorgelly P, Garellick G, Malchau H. The "hazards" of extrapolating survival curves. Med Decis Mak [Internet]. 2013 Apr [cited 2020 Aug 17];33(3):369–80. Available from: https://pubmed.ncbi.nlm.nih.gov/23457025/

10. Katkade VB, Sanders KN, Zou KH. Real world data: An opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. J Multidiscip Healthc. 2018;11:295–304.

11. Littlejohns P, Chalkidou K, Culyer AJ, Weale A, Rid A, Kieslich K, et al. National Institute for Health and Care Excellence, social values and healthcare priority setting [Internet]. Vol. 112, Journal of the Royal Society of Medicine. SAGE Publications Ltd; 2019 [cited 2021 May 11]. p. 173–9. Available from: https://journals.sagepub.com/doi/full/10.1177/0141076819842846

12. Schlander M. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Health Technology Assessments by the National Institute for Health and Clinical Excellence. Springer New York; 2007.

13. Broader types of data to be used in development of NICE guidance | News and features

| News | NICE [Internet]. [cited 2020 May 5]. Available from: https://www.nice.org.uk/news/article/broader-types-of-data-to-be-used-in-development-of-nice-guidance

14. Wang S, Goring SM, Lozano-Ortega G. Inclusion of real-world evidence in submission packages to health technology assessment bodies: What do current guidelines indicate? Value Heal [Internet]. 2016 May 1 [cited 2020 Jun 22];19(3):A287. Available from: http://www.valueinhealthjournal.com/article/S1098301516008317/fulltext

15. Pietri G, Masoura P. Market Access and Reimbursement: The Increasing Role of Real-World Evidence. Value Heal [Internet]. 2014 Nov 1 [cited 2020 Jun 22];17(7):A450–1. Available from: http://www.valueinhealthjournal.com/article/S1098301514031465/fulltext

16. Harwood M, Deighton A, Mickle A, Qian C, Szabo S. The use of real-world data in health technology assessment of medications for rare diseases. Value Heal [Internet]. 2019 Nov 1 [cited 2020 Sep 10];22:S863. Available from: http://www.valueinhealthjournal.com/article/S1098301519348296/fulltext

17. Roberts MH, Ferguson GT. Real-World Evidence: Bridging Gaps in Evidence to Guide Payer Decisions. PharmacoEconomics - Open [Internet]. 2021 Mar 18 [cited 2021 Jun 29];5(1):3–11. Available from: https://doi.org/10.1007/s41669-020-00221-y

18. Bullement A, Podkonjak T, Robinson MJ, Benson E, Selby R, Hatswell AJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. Int J Technol Assess Health Care [Internet]. 2020 Aug 1 [cited 2021 May 12];36(4):388–94. Available from: https://doi.org/10.1017/S0266462320000434

19. FDA. Real-World Evidence [Internet]. 2020 [cited 2020 May 6]. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

20. Guide to the multiple technology appraisal process [Internet]. [cited 2020 Aug 7]. Available from: www.nice.org.uk

Figure Legends

Figure 1 Inclusion/exclusion criteria

Figure 2 The framework for data extraction

Figure 3 Hypotheses about increased use of RWD

Supplement Legends

Supplement 1 Glossary of variables in extraction template

Inclusion criteria
- STA of oncology medicine
- Appraisals issued from January 2011 to May 2021
Exclusion criteria
- Appraisal of technology for preventing the complications of cancer
- Appraisal of surgical practice and other therapeutic therapies
- Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals)

Figure 1 Inclusion/exclusion criteria

131x35mm (240 x 240 DPI)

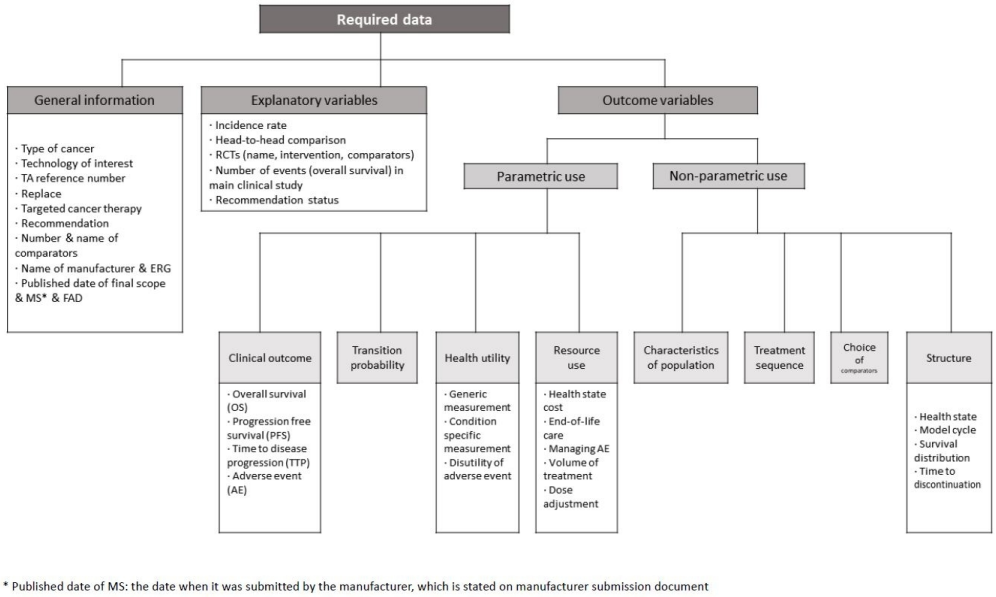


Figure 2 The framework for data extraction

141x84mm (240 x 240 DPI)

1) Poor internal/external validity of the clinical trial is associated with greater use of RWD.
2) Absence of direct (head-to-head) comparison is associated with greater use of RWD.
3) Low incidence rate of the disease is associated with greater use of RWD.
4) Immature survival data in the clinical trial are associated with greater use of RWD.
5) The technology having been recommended in previous NICE TA guidance is associated with greater use of RWD.

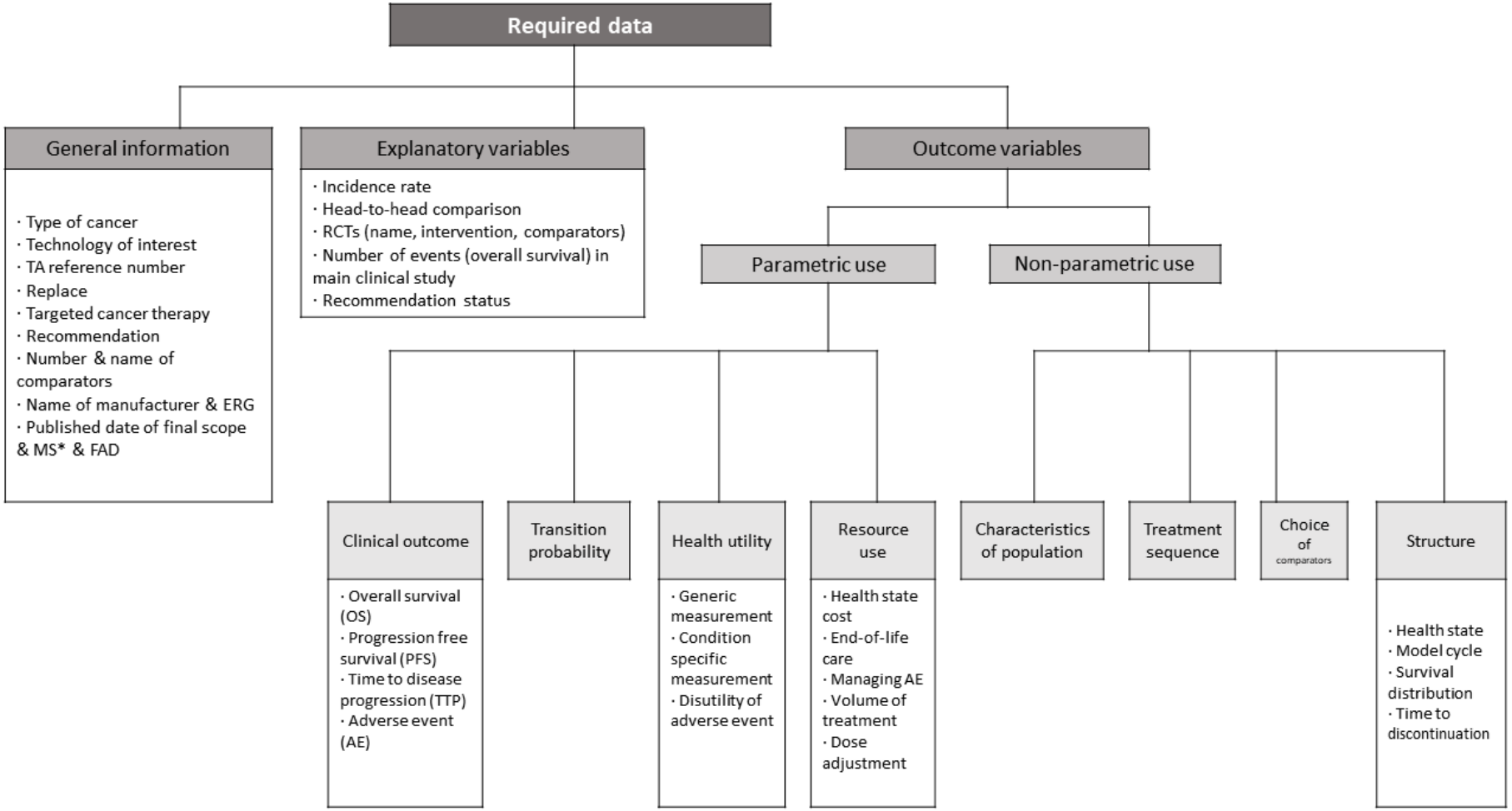
Figure 3 Hypotheses about increased use of RWD

93x24mm (240 x 240 DPI)

Figure 1 Inclusion/exclusion criteria

Inclusion criteria
- STA of oncology medicine
- Appraisals issued from January 2011 to May 2021
Exclusion criteria
- Appraisal of technology for preventing the complications of cancer
- Appraisal of surgical practice and other therapeutic therapies
- Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals)

Figure 2 The framework for data extraction



* Published date of MS: the date when it was submitted by the manufacturer, which is stated on manufacturer submission document

Figure 3 Hypotheses about increased use of RWD

1) Poor internal/external validity of the clinical trial is associated with greater use of RWD.
2) Absence of direct (head-to-head) comparison is associated with greater use of RWD.
3) Low incidence rate of the disease is associated with greater use of RWD.
4) Immature survival data in the clinical trial are associated with greater use of RWD.
5) The technology having been recommended in previous NICE TA guidance is associated with greater use of RWD.

Supplement 1 Glossary of variables in extraction template

General information		
Variable	**Explanation**	**Coding**
Type of cancer	The NICE classification of the cancer (website: https://www.nice.org.uk/guidance/conditions-and-diseases/cancer)	Bladder cancer=1, Blood and bone marrow cancer =2, Breast cancer=3, Colorectal=4, Neuroblastoma=5, Head and neck=6, Liver=7, Lung=8, Oesophageal=9, Ovarian=10, Pancreatic=11, Prostate=12, Renal=13, Skin=14, Stomach=15, Sarcoma=16
Technology of interest	The name of drug in the current appraisal. If it is combination therapy, the key technology which manufacturer focuses on will be taken here.	Narrative description
Indication	Clinical indications which are addressed in Final Appraisal Determination (FAD) document	Narrative description
TA number	the reference number of the technology guidance	Narrative description
Replace	Whether TA guidance has replaced or not. Appraisals can be replaced after rapid reviews/reviews/updates of previous appraisals or CDF reviews. Regardless of reasons of replacement, TA reference number which is replaced by this appraisal of interest will be recorded.	None= 0 If current appraisal replaces previous appraisal, the replaced TA reference number is recorded here.
• Pre-2016 CDF reconsideration	Before April 2016, the drug which was not reviewed or not recommended for routine commissioning by NICE can be used using the previous model of CDF. When new CDF was introduced in April 2016, these drugs in the old CDF were appraised by NICE to transit the model of CDF. This variable describe whether the appraisal of interest is an appraisal of the CDF reconsideration for the drug used in the old model of CDF before 2016.	No, it is not pre-2016 CDF reconsideration =0 Yes, it is a appraisal of pre-2016 CDF reconsideration =1
• 2016 CDF review	In April 2016, a new model of CDF was introduced. In the new model, an additional recommendation, recommended for use within the CDF is available when NICE appraising cancer drugs. The drug available via the CDF has to collect the data for further review for the routine commissioning after a certain period. As this mandated data collection can impact on the use of RWD, this variable allows to distinguish the appraisals, which RWD is more likely to be used.	No, it is not 2016 CDF review =0 Yes, it is 2016 CF review=1
Targeted cancer therapy	Treatment that uses drugs or other substances to identify and attack specific types of cancer cells	Non-targeted therapy = 0, targeted therapy = 1, not sure = Narrative description

Recommendation	the classification of recommendations made by the NICE committee in FAD document - Not recommended: 0 - Recommended (in line with marketing authorisation): 1 - Recommended (in line with marketing authorisation) in CDF:2 - Optimised: 3 - Optimised in CDF: 4 - Recommended in research: 5	Not recommended=0, recommended=1, recommended (cdf)=2, optimised=3, optimised (cdf)=4, recommended in research=5
number of comparators	Count the number of comparators in each manufacturer submission or FAD document. The information in manufacturer submission and FAD is recorded in the separated rows (manufacturer row/committee row).	Number in the manufacturer's submission
name of comparators	Record the name of comparators in manufacturer submission or FAD document	Narrative description
name of manufacturer	the name of manufacturer in manufacturer submission	Narrative description
name of the ERG	the name of the ERG (evidence review group)/AG (assessment group) in ERG critiques or AG reports	Narrative description
published date of final scope	the date of final scope as MM/YYYY	Date (MM/YYYY)
published date of manufacturer	the date of manufacturer submission as MM/YYYY.	Date (MM/YYYY)
published date of FAD guidance	the date of FAD document as MM/YYYY	Date (MM/YYYY)
Explanatory variables		
Variable	**Explanation**	**Coding**
Incidence (rate, year)	The rate would be recorded as it is in the appraisal. Incidence rate could be found in the final scope document or in manufacturer submission document. If the figures are not identical in each document, the latest rate is recorded. Most appraisals present the annual estimate of the number of patients who are eligible for the treatment in the "Budget Impact" section of company submission. This number is mainly used for the incidence. If this information is not available in the appraisal, the number in previous appraisal for similar indication is used instead.	Number
H2H	Whether the head-to-head clinical trial of a technology of interest exists or not, which compares with agreed comparators. The information is most likely to be found in the section: Identification and selection of relevant studies in clinical effectiveness part.	no=0, yes=1, yes but some comparators missing =2

• ITC	ITC (indirect treatment comparison). The information could be found in the section: Indirect and mixed treatment comparisons in clinical effectiveness part.	no=0, yes=1
• RCT (technology of interest)	Main RCT used in the appraisal: the name of the H2H RCT, if it exists. Unless there is an H2H, RCT refers to the clinical trial of technology of interest in the ITC.	no=0, yes=1
- Name of RCT	The name of the aforementioned RCT	Narrative description
- Intervention in RCT	The name of the intervention used in the aforementioned RCT. This variable helps to identify the main technology in RCT when technology is appraised as combination therapy.	Narrative description
- Comparators in RCT	The comparator of the aforementioned RCT	Narrative description
- Size of RCT	The number of participants in the aforementioned RCT	Number
- Median duration of follow-up	The median duration of follow-up in the aforementioned RCT. If it is not reported, record as NR (not reported).	Unit: month Not reported = ..
• Anchored/unanchored	"Anchored" means that RCT of technology of interest exists, and the RCT has been linked to any other studies which evaluate the drug's effectiveness. "Unanchored" means that the clinical outcome study doesn't have any comparators which connect to other studies. For example, comparing a single-arm study with a single-arm study is "unanchored". Also, RCTs compared without common comparators in ITC is "unanchored".	Not anchored=0, Anchored =1
• MAIC/STC	Matching adjusted indirect comparison (MAIC), Simulated Treatment Comparison (STC). A methodology of making adjustment to increase the comparability of two distinct populations mostly among unanchored studies. But it could be used in anchored studies in case where the two populations in ITC is starkly different from each other.	Naive=0, MAIC=1 STC=2 Other methods=3
Risk of bias (RoB) of RCT (direct quotation)	In order to evaluate the internal validity of RCTs, the risk of bias, which was reported in the ERG report, will be recorded here. Information is available at the quality assessment part of the ERG report. The ERG assesses the risk of bias of the included study using quality assessment tools. The ERG statement is directly quoted. The ERG often addresses the issue of quality of study narratively. Moreover, the ERG uses different terminology, whereas the domain of assessment is consistent. Therefore, the risk of bias would be narratively recorded. Prior to analysis, it will be scored by looking at the number of factors about which the ERG has expressed concern.	Direct quotation from ERG documents
• Risk of bias in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into four groups following the number of risk factors.	High/good quality without mentioned weakness= 0, risk factor 1 (low) =1, risk factor 2-3 (moderate)=2, risk factor 4

		(high) =3
External validity of RCT	As narrative accounts, generalisability of RCT is reported in the ERG report whether the population of RCT properly represents the UK general population in terms of aging structure, health status and health care practice (practice-dose, subsequent treatment, etc.).	Direct quotation from ERG documents
• External validity in RCT (grade)	In order to conduct statistical analysis, a set of codes will be used here. The direct quotation will be classified into three groups following the severity of generalisability assess by ERG.	Representative without mentioned weakness= 0, Representative but minor concerns =1, Questionable generalisability =2
Previously recommended in other indication	Whether the technology has been recommended for other types of cancers besides the current indication of the technology.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in other indication	If it was recommended for other indications, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Previous recommended treatment in the same cancer	Whether the technology has been recommended for other treatment lines in the same type of cancer.	No =0, Yes including all recommend, CDF, Optimised, Optimised (cdf) =1
• TA number & date of appraisal in the same cancer	If it was recommended for other treatment lines in the same cancer category, record the TA number and the date of the FAD documents (MM/YYYY).	Narrative description of date
Maturity of survival data in clinical trial	The data maturity is examined by looking at the number of events (deaths) of intervention arm in clinical trials. In published appraisal document, some of the information is redacted due to confidentiality. If the information is not available, the article of clinical trial published in journals is searched in order to check how many events are observed during the trial. Nonetheless, data are still not available in some cases. Since manufacturer is likely to redact the OS information when median OS was not reached. Hence, the survival data in this case are regarded as immature.	Direct quote from manufacturer submission
• Maturity (grade)	The direct quotation will be classified into three groups following the data cut point, 20% and 50 % of the number of events. This protocol adapts the criterion for measuring maturity of survival data in Tai et al. which investigates data maturity in STAs by looking at the proportion of death in pivotal trials. In the study, 20, 50 and 70 % of proportion of number of deaths are used to discuss the maturity of survival data (1). This protocol only uses 20% and 50% to assess the maturity without the category “unclear.”	Immature (number of events < 20%) =1, Relatively immature (20%≤number of events≤50%)=2 Mature (number of events < 50%) =3

Outcome variables

Variable	Explanation	Coding	Example

characteristic of population	Whether RWD are used to determine the characteristic of population, including the initiation age and health performance status (ECOG) or not. - Soft use: when RWD are supplementary evidence to decide the population characteristics - Hard use: when RWD determine the characteristics of population in economic evaluation	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Pomalidomide, in combination with low-dose dexamethasone, for treating multiple myeloma in adults at third or subsequent relapse (NICE TA427): baseline patient characteristics were obtained from RWD collected from a hospital population since the majority of the trial populations were previously untreated, which was different from target population.
treatment sequence	Whether RWD are used to determine the subsequent treatment option or not. After the disease progression onto the later stages of cancer treatments, patients are likely to receive idiosyncratic subsequent treatments. The pattern of subsequent treatment for cost-effectiveness analysis could be observed by RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): a study of medical records was used to determine the treatment sequence.
choice of comparator	Whether RWD are used to choose the comparators in economic evaluation or not. Although comparators are chosen based on the current clinical guideline, drug utilisation data or clinical expert opinion are frequently referred to find the most relevant comparators in evaluation.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Ixazomib with lenalidomide and dexamethasone for treating relapsed or refractory multiple myeloma (NICE TA505): the manufacturer considered that lenalidomide was appropriate comparator based on IMS market research data (lenalidomide, 69% market share and panobinostat, 7%).
structure (health state)	Whether RWD are used to determine the health state such as stable, progression, and death in a given model. Information is available at health state in the model of cost-effectiveness analysis in manufacturer submission documents.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Palbociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer (NICE TA495): the model health state of post-progression was specified based on a retrospective patient medical record review study.
structure (model cycle)	Whether RWD are used to determine model cycle or not. Model cycle, hereby, means that the duration between different health states, which can be influenced by the severity of conditions.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A **
Structure (survival distribution of intervention)	Whether RWD are used to decide the survival distribution of intervention or not. Since survival rate observed in RCTs is immature, it is necessary to extrapolate the survival rate for analysis. In order to choose proper survival distribution, the goodness of fit is tested (AIC, BIC).	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Larotrectinib for treating advanced solid tumours with NTRK fusions (NICE TA630): UK all-cause mortality data were used to assess the clinical acceptability of distributions whether patient overall survival exceeded current UK life expectancy

	Also, the clinical plausibility is asked to validate the distribution. In this case, the alternative data can be utilized. - If RWD is utilised for choosing distribution, mark as "hard use". - If RWD is utilised as supplementary evidence for the chosen distribution, mark as "soft use".		
Structure (survival distribution of comparator)	Whether RWD are used to validate the feasibility of survival distribution of comparator or not. As survival distributions of intervention and comparators are separately determined, the extraction tool approach it independently. Apply the abovementioned description on survival distribution of intervention to comparator in this row.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	
Structure (Time to discontinuation of intervention)	Whether RWD are used to decide the time to discontinuation of intervention or not. The time to discontinuation is likely to be decided by 1) simply adopting discontinuation rule in trials, 2) formulating distribution of discontinuation, or 3) clinical experts' opinion. - If RWD are used for designating the time to discontinuation, mark as "hard use" - If RWD are used as supplementary evidence for designating the time to discontinuation, mark as "soft use". - If clinical experts' opinions are used for designating the time to discontinuation, it is not regarded as RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Lorlatinib for previously treated ALK-positive advanced non-small-cell lung cancer (NICE TA628): The plausibility of the extrapolation of time on treatment was validated by UK RWD, hospital network data.
Structure (time to discontinuation of comparator)	Whether RWD are used to decide the time to discontinuation of comparator or not. Apply the above-mentioned description on time to discontinuation of intervention to comparator in this	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	

	row.		
Clinical outcome (OS) intervention	Whether RWD give the figure for overall survival (OS) of intervention or not. In order to measure the Quality Adjusted Life-Years (QALYs), it is necessary to extrapolate overall survival based on observed data on survival. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Nivolumab for adjuvant treatment of completely resected melanoma with lymph node involvement or metastatic disease (NICE TA558): the survival model applied the registry data (American Joint Committee on Cancer; AJCC) to both treatment arms after a certain time point.
Clinical outcome (PFS) intervention	Whether RWD give the figure for progression free survival (PFS) of intervention or not. The progression of disease is important for economic evaluation model in terms of health state transitions and treatment switching. The survival data could come from RCT or RWD.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (RR) intervention	Whether RWD provides the response rate (RR) for the intervention or not. The effectiveness of cancer treatment is often shown by responses of tumour cells, which is evaluated by the RECIST criteria or other criteria. The response rate data would be collected in RCT or other type of data.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (TTP) intervention	Whether RWD give the figure for time-to-progression (TTP) of intervention or not. Some cancer treatments show their clinical effectiveness not through the progression free survival (PFS), but alternatively through time-to-progression.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (AE) intervention	Whether RWD give the figure of adverse event (AE) of intervention or not. Adverse events are crucial information for the estimation of the QALYs. The adverse events are collected in RCT. However, RWD, including cohort studies, retrospective studies, or other type of studies, also provide the information of adverse events, which cannot be found in RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Blinatumomab for treating acute lymphoblastic leukaemia in remission with minimal residual disease activity (NICE TA589): retrospective non-interventional cohort study collected from 2000 to 2017 was used to inform the clinical outcome of comparators as well as adverse event.
Clinical outcome (OS) comparators	Whether RWD give the figure of overall survival (OS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, clinical outcome (OS) intervention

Clinical outcome (PFS) comparators	Whether RWD give the figure for the progression free survival (PFS) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (RR) comparators	Whether RWD provide the response rate (RR) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (TTP) comparators	Whether RWD provide the time-to-progression (TTP) of comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Clinical outcome (AE) comparators	Whether RWD provide the figure adverse events (AE) for the comparators or not.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, clinical outcome (AE) intervention
Transition probability	Whether RWD provide the transition probability from one state to other state, if it is applicable.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Pembrolizumab for treating melanoma with high risk of recurrence (NICE TA553): electronic health records (Flatiron database) collected by cancer care providers in the US was used to model transition from the "locoregional recurrence (LR)" state to the "distant metastases" and life tables for transition from the LR to "death" state.
Health utility of health state (generic)	Whether health state utility survey of generic measurement is done in RWD or RCT. Health state utility is necessary information for the estimation of the QALYs. Generic health utility measurement, EQ-5D, is frequently used. There is national tariff of EQ-5D to get the scores. Hereby, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Health utility of health state (condition-specific)	Whether health state utility survey of condition-specific measurement is done in RWD or RCT. In cancer treatment, condition-specific measurement is commonly adopted. Similar to the previous row, the way of collecting survey (RWD or RCT) is highlighted.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**

Disutility of adverse events	Whether survey of collecting disutility data is done in RWD or RCT. As adverse events are likely to reduce the patient's quality of life, the disutility of adverse events is included in estimates. The way of collecting survey (RWD or RCT) is drawn to attention.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Health state cost) common	Whether resource use for estimating health state cost is derived from RWD or RCT. In economic evaluation, the unit cost mostly comes from the national reference cost. The total cost is calculated by the total resource use (volume of technology and health care services) multiplied by the reference cost. Here, the only resource use is focused in data extraction. Resource use for estimating health state cost includes all activity like monitoring, GP visits, pharmacy cost etc. Health state resource use could be aggregated or individually listed. Here, the difference of describing health state cost is not separately considered.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies (NICE TA559): RWD was used for estimating the cost of inpatient admission (data: Hospital Episode Statistics), the cost of home care and hospice (data: National Audit Office), and GP time (data: Personal Social Services Research Unit; PSSRU).
Resource use (end-of-life care) common	Whether resource use for estimating end-of-life care is derived from RWD or RCT. Resource use of terminal cancer patients is not frequently reported in the RCT providing the treatment effect. Therefore, other data resources, including RCTs of other technologies, provide the information of resource use in the end-of-life care.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Managing AE) intervention	Whether resource use for managing adverse events of intervention is derived from RWD or RCT. Resource use of managing adverse events is reported in RCTs as well as in other types of researches which can provide alternative perspectives.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (volume of treatment) intervention	Whether resource use for volume of treatment of intervention is derived from RWD or RCT. In this study, scope of the volume of treatment is limited to the frequency of treatment, frequency of administration, and amount of subsequent treatment.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	- Fulvestrant for treating untreated locally advanced or metastatic oestrogen-receptor positive breast cancer (NICE TA503): a medical chart review study was used to determine the proportion of patient using subsequent treatment for cost calculation.

Resource use (Dose adjustment) intervention	Whether resource use for dose adjustment of intervention is derived from RWD or RCT. There are several reasons for adjusting dose such as adverse events (AEs). The dose of cancer treatments is calculated by BSA (body surface area). This study focuses only on BSA and dose adjustment due to AEs, because these information are commonly reported in NICE appraisals.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (Managing AE) comparators	Whether resource use for managing adverse events of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
Resource use (volume of treatment) comparators	Whether resource use for volume of treatment of comparators is derived from RWD or RCT.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	Refer to the variable, resource use (volume of treatment) intervention
Resource use (Dose adjustment) comparators	Whether resource use for dose adjustment of comparators is derived from RWD or RCT. Since the intervention is a novel technology, RCTs provide less information on the adjustment. RWD could be utilised to provide more relevant information regarding dose adjustment of existing technologies which have been used in routine clinical practice.	No RWD = 0 Yes, data from RWD = 1 Not clear = 9	N/A**
* In order to detect the use of RWD in sensitivity analysis, the parametric part is duplicated. ** As data extraction is not conducted, all of examples are not available at this stage. In this case, it marked as N/A. *** Benefits/challenges of the use of RWD are collected in outcome variables. **** In cases where trials have more than two arms, only the arms considered as relevant for decision problem in evidence submission are included. If there are two intervention arms and these arms are separately used for different indications in appraisals, the data extraction is carried out separately. When two arms are relevant as comparators for same indication, the data are recorded without distinguishing these arms.			

1. Tai TA, Latimer NR, Benedict A, Kiss Z, Nikolaou A. Prevalence of Immature Survival Data for Anti-Cancer Drugs Presented to the National Institute for Health and Care Excellence and Impact on Decision Making. Value Heal. 2020 Dec 8;