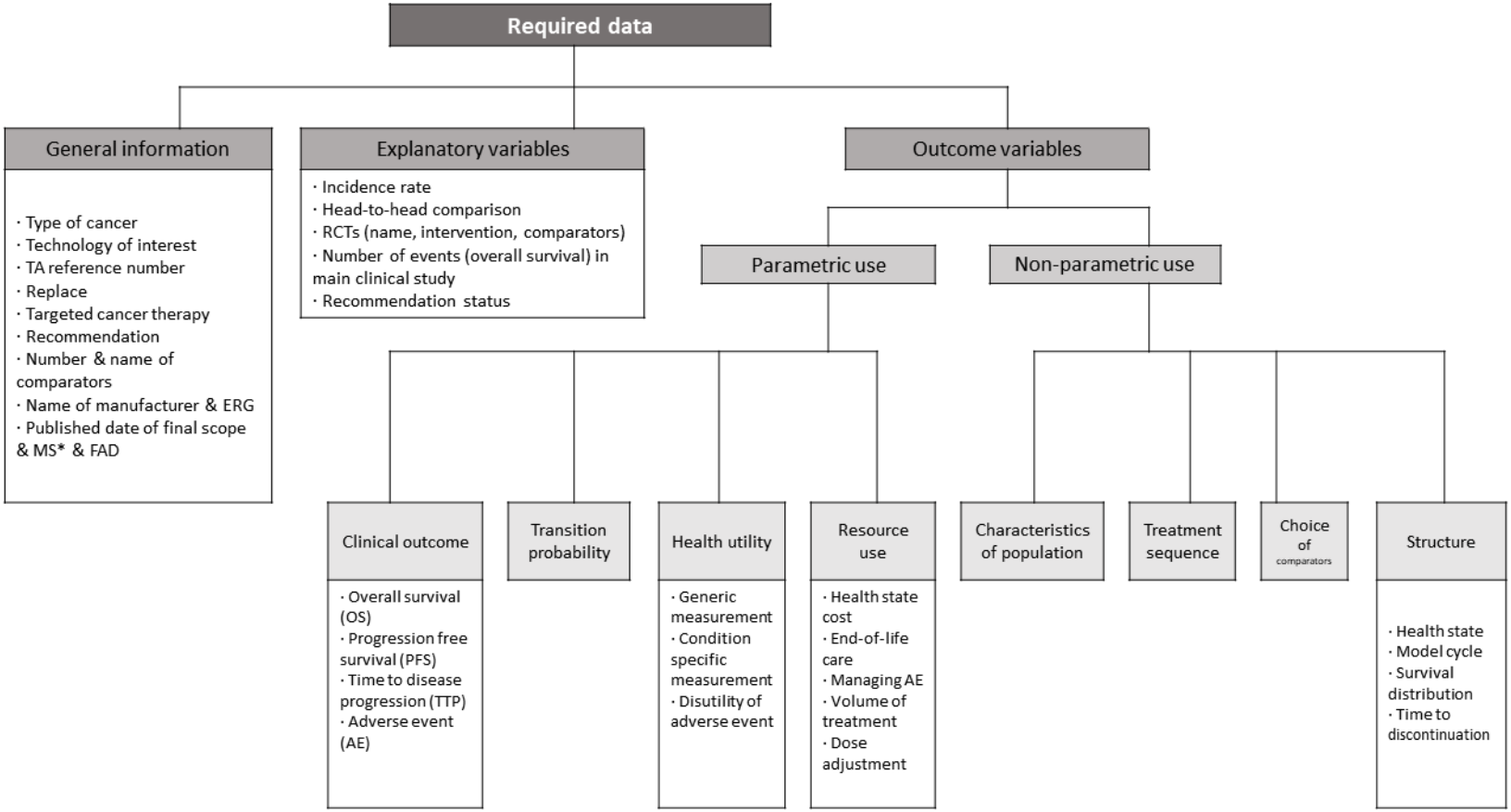
Figure 1 Inclusion/exclusion criteria

| |
|---|
| 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| - Appraisals for which evidence is not available (withdrawn appraisals) or was never supplied (terminated appraisals) |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<br>- Not recommended: 0<br>- Recommended (in line with marketing authorisation): 1<br>- Recommended (in line with marketing authorisation) in CDF:2<br>- Optimised: 3<br>- Optimised in CDF: 4<br>- 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.<br>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<br>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.<br>"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,<br>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,<br>MAIC=1<br>STC=2<br>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.<br>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,<br>Relatively immature (20%≤number of events≤50%)=2<br>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.<br>- Soft use: when RWD are supplementary evidence to decide the population characteristics<br>- Hard use: when RWD determine the characteristics of population in economic evaluation | No RWD = 0<br>Yes, data from RWD = 1<br>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.<br>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<br>Yes, data from RWD = 1<br>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.<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>Not clear = 9 | N/A ** |
| Structure (survival distribution of intervention) | Whether RWD are used to decide the survival distribution of intervention or not.<br><br>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<br>Yes, data from RWD = 1<br>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.<br><br>- If RWD is utilised for choosing distribution, mark as "hard use".<br><br>- If RWD is utilised as supplementary evidence for the chosen distribution, mark as "soft use". | | |
| Structure<br>(survival distribution of comparator) | Whether RWD are used to validate the feasibility of survival distribution of comparator or not.<br><br>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<br>Yes, data from RWD = 1<br>Not clear = 9 | |
| Structure<br>(Time to discontinuation of intervention) | Whether RWD are used to decide the time to discontinuation of intervention or not.<br><br>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.<br>- If RWD are used for designating the time to discontinuation, mark as "hard use"<br>- If RWD are used as supplementary evidence for designating the time to discontinuation, mark as "soft use".<br>- If clinical experts' opinions are used for designating the time to discontinuation, it is not regarded as RWD. | No RWD = 0<br>Yes, data from RWD = 1<br>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<br>(time to discontinuation of comparator) | Whether RWD are used to decide the time to discontinuation of comparator or not.<br><br>Apply the above-mentioned description on time to discontinuation of intervention to comparator in this | No RWD = 0<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>Not clear = 9 | N/A** |
| Clinical outcome (RR) comparators | Whether RWD provide the response rate (RR) of comparators or not. | No RWD = 0<br>Yes, data from RWD = 1<br>Not clear = 9 | N/A** |
| Clinical outcome (TTP) comparators | Whether RWD provide the time-to-progression (TTP) of comparators or not. | No RWD = 0<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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<br>Yes, data from RWD = 1<br>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.<br>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<br>Yes, data from RWD = 1<br>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;