European Journal of Cancer 209 (2024) 114262

Contents lists available at ScienceDirect

European Journal of Cancer

journal homepage: www.ejcancer.com

Using non-randomized trials to assess the clinical benefit of systemic anti-cancer treatments: Viable or not?

N.S.H. Xander [a,b,*], B. Leeneman [a,b], A.-M.C. Dingemans [c], W.E. Fiets [d], W.K. de Jong [e], N.E.M. Uyl [a], A.N.M. Wymenga [f], A.K.L. Reyners [g], C.A. Uyl-de Groot [a,b]

[a] Department of Health Technology Assessment, Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands
[b] Erasmus Centre for Health Economics Rotterdam, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands
[c] Department of Pulmonary Medicine, Erasmus MC Cancer Institute, Erasmus Medical Center, Dr. Molewaterplein 40, 3015 GD Rotterdam, the Netherlands
[d] Department of Medical Oncology, Medical Center Leeuwarden, Henri Dunantweg 2, 8934 AD Leeuwarden, the Netherlands
[e] Department of Pulmonology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands
[f] Department of Medical Oncology, Medisch Spectrum Twente, Koningsplein 1, 7512 KZ Enschede, the Netherlands
[g] Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, the Netherlands

ARTICLE INFO

Keywords:
Randomized controlled trial
Non-randomized trial
Clinical value assessment
Comparison
Cancer

ABSTRACT

Background: The Dutch Committee for the Evaluation of Oncological Agents (cieBOM) assesses the clinical benefit of systemic anti-cancer treatments (SACTs). For SACTs tested in non-randomized trials (NRTs), cieBOM primarily utilizes response-related thresholds as assessment criteria. As sufficiency of NRT-based evidence for benefit assessments is questionable, this study investigated whether and how NRTs can be used to assess the clinical benefit of new SACTs initially appraised by cieBOM based on randomized controlled trials (RCTs).

Methods: Using the RCTs underpinning cieBOM recommendations issued between 2015 and 2017, we searched for matching NRTs and applied the NRT-related assessment criteria by cieBOM to them. We then compared the assessment outcomes to the respective RCT-based cieBOM recommendations. Further, we investigated how the assessments would change when applying different response-related thresholds and adding a progression-free survival (PFS) threshold.

Results: For 13 of the 37 eligible recommendations, a matching NRT was found. Two treatments were assessed positively and six negatively; five treatments were non-assessable. Two positive recommendations matched a positive NRT-based assessment; one matching negative assessment was found, and one treatment could not be assessed based on either trial results. Adding a $>$ 6 months PFS threshold decreased the number of non-assessable NRTs (five to two).

Conclusions: Limited publications and inconsistent data reporting hampered the viability of NRTs for clinical benefit assessments of SACTs beyond the scope of rare indications. Further, response-related assessment criteria alone might not fully grasp the clinical benefit of novel SACTs. NRT-based assessments should be considered with caution due to uncertainty of the trial results.

1. Introduction

Randomized controlled trials (RCTs) are viewed as the gold standard for assessing the efficacy and safety of new systemic anti-cancer treatments (SACTs). Therefore, the European Medicines Agency (EMA) generally uses them as a basis for their assessment [1–3]. However, to facilitate early patient access to new treatments in rare indications, EMA increasingly bases its assessments on non-randomized trials (NRTs) [4–7].

After a new SACT is granted market authorization, the magnitude of its clinical benefit is assessed by the European Society for Medical Oncology (ESMO) using the Magnitude of Clinical Benefit Scale (MCBS) [8,9]. Originally developed in 2015 to assess treatments based on RCT results [10], the revised version of this tool facilitates the grading of

* Correspondence to: Department of Health Technology Assessment, Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, the Netherlands.
E-mail address: xander@eshpm.eur.nl (N.S.H. Xander).

https://doi.org/10.1016/j.ejca.2024.114262

Available online 3 August 2024

treatments tested in NRTs [11]. The crucial primary endpoints of such studies should be objective response rate (ORR) with or without duration of response (DoR) or progression-free survival (PFS). Additionally, quality of life (QoL) improvement and the prevalence of grade 3/4 adverse events (AEs) also factor in with the score [11].

In the Netherlands, the Committee for the Evaluation of Oncological Agents (cieBOM; *Commissie ter Beoordeling van Oncologische Middelen* in Dutch) assesses the clinical benefit of newly registered SACTs. The assessments are based on the so-called PASKWIL criteria. These criteria, which were established in 1999 and are regularly updated [8,9], involve effectiveness, treatment burden, AEs and QoL [12]. By applying these criteria, cieBOM provides recommendations on whether a new treatment should be considered for use within Dutch clinical practice [8].

In 2021, cieBOM introduced assessment criteria suitable for NRTs (NRT-PASKWIL criteria) [13]. Currently, they are only applied to SACTs for rare indications and generally involve ORR and DoR. PFS and OS are only considered for the assessment if there is an implicit or explicit control group in the relevant NRT. AEs are disregarded (see Table 1).

Since NRT results are less robust than those obtained in an RCT, it is questionable to what extent an NRT can provide sufficient evidence for a conclusive clinical benefit assessment of SACTs. Therefore, by applying the NRT-PASKWIL criteria, we investigated whether and how NRTs can be used to assess the clinical benefit of new SACTs initially appraised by cieBOM based on RCTs.

2. Methods

2.1. Study selection

CieBOM's assessment reports are published on the website of the Dutch Society for Medical Oncology (NVMO) [14]. From this website, all recommendations on non-curative SACTs published from January 2015 to December 2017 were selected. Reports concerning adjuvant therapies, reassessments and recommendations based on phase II RCTs were excluded.

To identify the available relevant NRTs, we implemented a matching procedure with several underlying conditions. We conducted a review of potential matches to the phase III RCTs underpinning the recommendations by cieBOM as well as a narrative review in Google Scholar, PubMed, and ClinicalTrials.gov using the treatment combined with "phase I", "phase II", "non-randomized", and "single-arm" as search terms. Matching criteria included treatment, patient population, and treatment line. The NRT had to precede the RCT, with non-overlapping enrolment periods, unless the NRT was part of the same study trajectory or specifically referenced in the RCT-related publication, with a maximum overlap period of 6 months. If the NRT enrolment period was not (fully) reported, the result publication date was assumed to indicate precedence to the matching RCT. NRT results had to be published in a peer-reviewed journal before the results of the RCT and report at least one outcome of interest (ORR, DoR, or PFS). Results reported in conference abstracts were excluded.

Multi-arm trials assessing a specific treatment across various drug dosages or treatment lines, lacking a comparator such as placebo or standard treatment, were considered NRTs. Only the treatment arm aligned with EMA-approved dosage and treatment line was pertinent for matching. The same applied to studies that randomized patients into groups based on drug dosage or treatment line.

2.2. Data analysis

Data were collected from cieBOM's assessment reports, the underpinning phase III RCT results, and ESMO-MCBS scorecards, and processed in a standardized data collection form in Microsoft Excel®. We extracted: report details (date of publication on the NVMO website, treatment, and indication); study characteristics of the underlying RCT and the available matching NRT (indication, study phase, treatment arm [s], enrolment period, and primary endpoint[s]); reported outcomes (ORR and/or DoR, PFS estimates, AE prevalence).

Secondly, the NRT-PASKWIL criteria were applied to assess the clinical benefit of the SACTs tested in the matching NRTs. Table 1 contains an overview of these criteria. A positive assessment required that the applicable ORR and DoR combination criterion was fulfilled. If neither or only one of the thresholds was exceeded, the assessment was negative. In case of incomplete reporting (e.g., no ORR confidence interval or DoR reported), the treatment was considered not assessable. The OS/PFS criterion as per the NRT-PASKWIL criteria was not applied due to a lack of control groups in the available NRTs.

Moreover, several alternative criteria were formulated to determine how the overall assessments would change upon their application (see Table 2). For alternatives 1–3, if at least the ORR or DoR threshold was met, the assessment was scored as positive. If none of the ORR or DoR values were sufficiently reported, the treatment was considered not assessable. In alternative 4, a $\geq$ 60% ORR threshold was applied without

Table 2
Current and alternative criteria for NRT-based assessments.

	Alternative criteria
Currently applied NRT-PASKWIL criteria	ORR > 40% (lower CI bound) AND DoR > 4 months; OR ORR > 30% (lower CI bound) AND DoR > 8 months; OR ORR > 20% (lower CI bound) AND DoR > 12 months
Alternative 1	ORR > 40% (lower CI bound) OR DoR > 4 months
Alternative 2	ORR > 30% (lower CI bound) OR DoR > 8 months
Alternative 3	ORR > 20% (lower CI bound) OR DoR > 12 months
Alternative 4	Median ORR $\geq$ 60%, no DoR
Alternative 5	Median PFS > 6 months (considered for assessment if PFS is primary endpoint of trial and ORR and DoR data were reported insufficiently)
Alternative 6	Median PFS > 6 months (always considered if ORR and DoR data were reported insufficiently)
Addendum	ESMO-MCBS score attributed to each study

Abbreviations: CI, confidence interval; DoR, duration of response; ESMO-MCBS, European Society for Medical Oncology – Magnitude of Clinical Benefit Scale; NRT, non-randomized trial; ORR, objective response rate; PFS, progression-free survival.

Table 1
PASKWIL criteria applied by CieBOM to non-randomised studies (NRT-PASKWIL criteria 2021) [8].

- The indication for which the treatment is registered is rare.
- For the indication, there are no treatment options for which a clinical advantage has been determined, or such treatment options have been exhausted.
- The population regarding the indication must be selected reliably and reproducibly.
- Preferably, there is a biological rationale for the use of the treatment.
- In considering the clinical value based on the objective response rate (ORR), a combination criterion is used with the lower bound of the 95% confidence interval (95% CI) of the ORR and the point estimate of the median duration of response (DoR). The researched treatment is assessed as clinically relevant and provides added value for the target group if:
  - ORR > 40% and DoR > 4 months or
  - ORR > 30%–40% and DoR > 8 months or
  - ORR > 20%–30% and DoR > 12 months
- In a non-randomized study, overall survival (OS) or progression-free survival (PFS) has added clinical value if the gain in OS or PFS compared to the implicit or explicit control group amounts to more than 16 weeks (lower bound 95% CI).

any DoR threshold. Alternatives 5 and 6 included a median PFS of > 6 months as an additional criterion with specific prerequisites: if PFS was the primary endpoint of the study, and no sufficient data was reported on ORR and DoR, then the appraisal of PFS was used for a preliminary assessment of the treatment based on the NRT-PASKWIL criteria (alternative 5). As a different alternative, PFS was considered for a preliminary assessment in case of insufficient reporting on ORR and DoR data regardless of the primary endpoint of the study (alternative 6). Further, a tentative ESMO-MCBS score was attributed to each treatment tested in the included NRTs based on the ESMO-MCBS Evaluation Form 3, with a higher score (the highest being 4) indicating a greater clinical benefit [11].

Finally, the outcomes of the assessments on the SACTs based on the current and the alternative criteria were compared to the cieBOM recommendations for the respective matching RCT, which were based on the applicable PASKWIL-criteria (RCT-PASKWIL criteria).

3. Results

3.1. Availability of matching NRTs

Overall, 44 recommendations by cieBOM issued between January 2015 and December 2017 were identified in the search. After removing recommendations on adjuvant treatments (n = 1), reassessments (n = 2) and randomized phase II studies (n = 4), 37 recommendations remained eligible. Hereof, 36 were based on phase III RCTs, and one on a phase II/III RCT.

For 21 of the remaining 37 relevant RCTs, no adequate match was available. In six cases, no matching published NRT was found. Eight studies had a mismatch in at least one relevant criterion (e.g., patient group, treatment line). Four cases saw an excessive overlap period (n = 2) or a delayed publication of matching NRT results (n = 2). Three phase II studies were designed as an RCT and therefore ineligible. Results of three matching studies were only published as an abstract. The remaining 13 studies were either designed as an NRT or considered as such since the randomization concerned the same treatment [15–27]. Figure 1 summarizes the results of the selection procedure; Supplementary Tables 1 and 2 provide a detailed overview of the relevant studies and matching trial characteristics, respectively.

3.2. Applicability of current NRT-PASKWIL criteria and alternative criteria

Regarding the SACTs tested in the RCTs that were assessed by cieBOM between 2015 and 2017 and for which matching NRTs were found, 10 of the 13 recommendations (77%) were positive and two (15%) were negative; one relevant SACT (8%) could not be assessed (see Table 3). By contrast, using the NRT-PASKWIL criteria, two of the 13 included NRTs (15%) were assessed positively and six (46%) negatively, as is shown in Table 3. For five studies (38%), no assessment was possible. In six cases, the median DoR was either not reached or not reported; in eight cases, the ORR was reported without a CI. However, a negative assessment could still be given in three cases where the median ORR was below 20%.

Applying the alternative assessment criteria to NRT results revealed disparities regarding the outcomes. Alternatives 1–3 showed an increase in positive assessments (7/4/4, respectively) and a decrease in negative assessments (0/3/4, respectively) compared to the current criteria (2 positive, 6 negative). However, the number of studies remaining non-assessable based on ORR and DoR prerequisites showed negligible changes. Using a ≥ 60% ORR threshold and eliminating the DoR prerequisite (alternative 4) allowed assessments of all NRTs. However, while previously negative assessments remained negative, only one previously non-assessable study was assessed positively, four others negatively.

Adding the PFS threshold under the application prerequisite of PFS

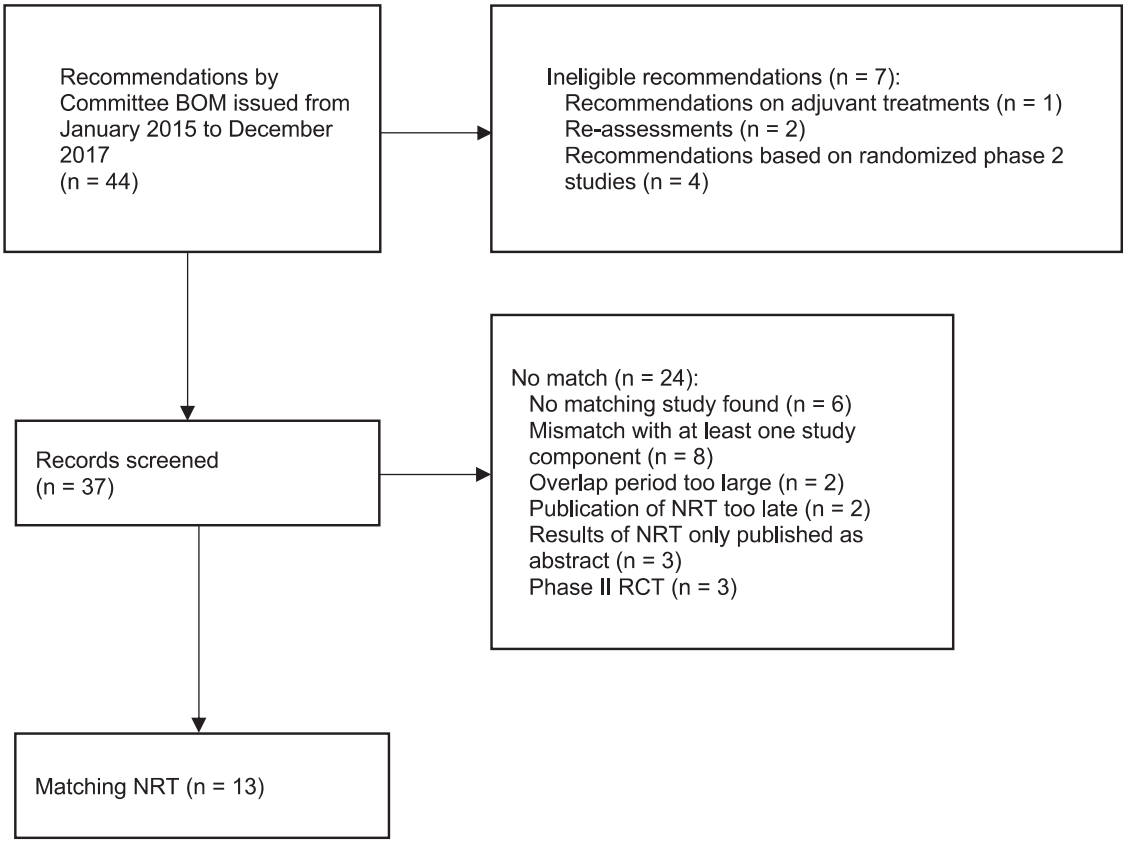

Fig. 1. Flowchart of matching results of cieBOM recommendations based on phase 3 RCTs and NRTs for the same treatment.

Table 3
Non-randomized trials included for analysis, assessment of included studies based on RCT-PASKWIL criteria and NRT-PASKWIL criteria.

Treatment	Indication	RCT: First Author, year of publication	NRT Phase	NRT: First author, year of publication	Primary endpoint (s)	ORR (95% CI)	Median DoR (months)	RCT-based recommendation by cieBOM	Assessment based on NRT-PASKWIL criteria
Cabozantinib	Advanced medullary thyroid cancer	Elisei, 2013 [28]	1[a]	Kurzrock, 2011 [15]	Safety	29% (15 −45%)	Not reached (range: 4 −35)	Positive	Negative
Bevacizumab	Platinum-resistant ovarian cancer	Pujade-Lauraine, 2014 [29]	2[b]	McGonigle, 2011 [16]	PFS	25% (NR)	NR	Positive	Not assessable
Bevacizumab	Advanced cervical cancer	Tewari, 2014 [30]	2[b]	Monk, 2009 [17]	PFS; toxicity	11% (NR)	6.21	Positive	Negative
Pembrolizumab	Advanced melanoma	Robert, 2015 [31]	1[b]	Hamid, 2013 [18]	Safety	38% (25 −44%)	Not reached (range: 1.9 −10.8)	Positive	Negative
Vemurafenib+cobimetinib	Non-resectable/ metastasized, BRAF-mutated melanoma	Larkin, 2014 [32]	1b[c]	Ribas, 2014 [19]	Safety; dose-limiting toxic effects; maximum tolerated dose	87% (NR)	12.5	Positive	Not assessable
Lenvatinib	Progressive, refractory thyroid cancer	Schlumberger, 2015 [33]	2[b]	Cabanillas, 2015 [20]	ORR	50% (37 −63%)	12.7	Positive	Positive
Nintedanib+docetaxel	Advanced/ metastasized non-small-cell lung cancer	Reck, 2014 [34]	2[d]	Reck, 2011 [21]	PFS; ORR	46% (NR)	NR	Negative	Not assessable
Nivolumab	Advanced/ metastasized clear-cell renal cell carcinoma (2nd/3rd-line treatment)	Motzer, 2015 [35]	2[d]	Motzer, 2015 [22]	PFS	20% (80% CI: 13.4 −29.1%)	22.3	Positive	Negative
Cabozantinib	Advanced clear-cell renal cell carcinoma (2nd-line)	Choueiri, 2015 [36,37]	1[b]	Choueiri, 2014 [23]	Safety; tolerability	28% (NR)	NR	Positive	Not assessable
Eribulin	Metastasized breast carcinoma (2nd-line)	Kaufman, 2015 [38]	2[b]	Vahdat, 2009 [24]	ORR	13.6% (NR)	5.6	Negative	Negative
T-VEC	Advanced, non-resectable melanoma	Andtbacka, 2015 [39]	2[b]	Senzer, 2009 [25]	Response rate	26% (NR)	NR	Not assessable	Not assessable
Osimertinib	Non-small-cell lung cancer with EGFR-T790M mutation	Mok, 2016 [40]	2[b]	Goss, 2016 [26]	Safety	70% (64 −77%)	11.4	Positive	Positive
Regorafenib	Hepatocellular carcinoma (2nd-line)	Bruix, 2017 [41]	2[b]	Bruix, 2013 [27]	PFS	3% (NR)	5.5	Positive	Negative

Abbreviations: CI, Confidence interval; cieBOM, Committee for the Evaluation of Oncological Agents; DoR, duration of response; EGFR, epidermal growth factor receptor; NR, not reported; NRT, non-randomized trial; ORR, objective response rate; PFS, progression-free survival; RCT, randomized controlled trial; T-VEC, talimogene laherparepvec.

[a] Dose-escalation study;
[b] Single-arm trial;
[c] Distinction between patients who recently progressed on a BRAF inhibitor and BRAF-inhibitor-naïve patients, incl. dose-escalation phase;
[d] Randomization concerns drug dosage.

being a primary endpoint (alternative 5) left two treatments (15%) not assessable based on their respective NRTs. Two treatments (15%) were assessed positively and six (46%) negatively. Two interventions (15%) were given a preliminarily positive assessment, and one (8%) a preliminarily negative assessment. When PFS was added as a criterion regardless of primary trial endpoints (alternative 6), one treatment (8%) remained not assessable. The now assessable treatment received a preliminarily positive assessment, resulting in three trials (23%) being likewise assessed. The remaining assessments remained unaffected compared to alternative 5.

Attributing an ESMO-MCBS score to each treatment based on reported NRT results showed that three treatments would receive a score of 1, four score 2, and five score 3 (see Table 4). One treatment (Talimogene laherparepvec; T-VEC) could not be graded due to a lack of

Table 4
Assessments of included NRTs based on criteria applied in each of the scenarios; tentative ESMO-MCBS score based on the criteria in Form 3 [11].

Treatment	**Indication**	**RCT-based assessment by cieBOM**	**Assessment based on NRT-PASKWIL criteria**	Alternative **1: ORR > 40% (lower CI bound) OR DoR > 4 months**	Alternative **2: ORR > 30% (lower CI bound) OR DoR > 8 months**	Alternative **3: ORR > 20% (lower CI bound) OR DoR > 12 months**	Alternative **4: Median ORR ≥ 60%**	Alternative **5: Median PFS > 6 months (applied if PFS was primary endpoint and ORR and DoR were reported insufficiently)**	Alternative **6: Median PFS > 6 months (criterion applied regardless of endpoint in the trial if ORR and DoR were reported insufficiently)**	Addendum: **ESMO-MCBS score attributed to each study**
Cabozantinib	Advanced medullary thyroid cancer	Positive	Negative	Not assessable	Not assessable	Not assessable	Negative	Negative	Negative	1
Bevacizumab	Platinum-resistant ovarian cancer	Positive	Not assessable	Not assessable	Not assessable	Not assessable	Negative	Preliminarily positive	Preliminarily positive	3
Bevacizumab	Advanced cervical cancer	Positive	Negative	Positive	Negative	Negative	Negative	Negative	Negative	2
Pembrolizumab	Advanced melanoma	Positive	Negative	Not assessable	Not assessable	Negative	Negative	Negative	Negative	3
Vemurafenib+cobimetinib	Non-resectable/ metastasized, BRAF-mutated melanoma	Positive	Not assessable	Positive	Positive	Positive	Positive	Not assessable	Preliminarily positive	2
Lenvatinib	Progressive, refractory thyroid cancer	Positive	Positive	Positive	Positive	Positive	Negative	Positive	Positive	2
Nintedanib+docetaxel	Advanced/ metastasized non-small-cell lung cancer	Negative	Not assessable	Not assessable	Not assessable	Not assessable	Negative	Preliminarily negative	Preliminarily negative	2
Nivolumab	Advanced/ metastasized clear-cell renal cell carcinoma (2nd/ 3rd-line treatment)	Positive	Negative	Positive	Positive	Positive	Negative	Negative	Negative	3
Cabozantinib	Advanced clear-cell renal cell carcinoma (2nd-line)	Positive	Not assessable	Not assessable	Not assessable	Not assessable	Negative	Preliminarily positive	Preliminarily positive	3
Eribulin	Metastasized breast carcinoma (2nd-line)	Negative	Negative	Positive	Negative	Negative	Negative	Negative	Negative	1
T-VEC	Advanced, non-resectable melanoma	Not assessable	Not assessable	Not assessable	Not assessable	Not assessable	Negative	Not assessable	Not assessable	Not assessable
Osimertinib	Non-small-cell lung cancer with EGFR-T790M mutation	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	4
Regorafenib	Hepatocellular carcinoma (2nd-line)	Positive	Negative	Positive	Negative	Negative	Negative	Negative	Negative	1
Assessment		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Positive		10 (77%)	2 (15%)	7 (54%)	4 (31%)	4 (31%)	2 (15%)	2 (15%)	2 (15%)	
Preliminarily positive								2 (15%)	3 (23%)	
Negative		2 (15%)	6 (46%)	0	3 (23%)	4 (31%)	11 (85%)	6 (46%)	6 (46%)	
Preliminarily negative								1 (8%)	1 (8%)	
Not assessable		1 (8%)	5 (38%)	6 (46%)	6 (46%)	5 (38%)	0	2 (15%)	1 (8%)	

Abbreviations: CI, confidence interval; cieBOM, Committee for the Evaluation of Oncological Agents; DoR, duration of response; EGFR, Epidermal growth factor receptor; ESMO-MCBS, European Society of Medical Oncology – Magnitude of Clinical Benefit Scale; NA, not available; NRT, non-randomized trial; ORR, objective response ratio; PFS, progression-free survival; RCT, randomized controlled trial; T-VEC, talimogene laherparepvec.

necessary data [25]. Two of the included NRTs involved a QoL analysis [21,26]; one treatment, osimertinib for non-small-cell lung cancer (NSCLC) with *EGFR*-T790M mutation [26], received the highest possible score (4). Nivolumab for advanced clear-cell renal cell carcinoma (accRCC) received an ESMO-MCBS score of 3 but was negatively assessed under the current NRT-PASKWIL-criteria and most alternative criteria [22]. Similarly, pembrolizumab for advanced melanoma was negatively assessed under alternatives 5 and 6 [18]. Osimertinib for NSCLC with *EGFR*-T790M mutation was consistently assessed positively throughout all alternatives [26]. Bevacizumab (for platinum-resistant ovarian cancer) and cabozantinib (as second-line treatment for accRCC) were mostly not assessable, but positively assessed wherever the median PFS (estimated at $>$ 6 months in both NRTs [23,26]) was an applicable criterion.

3.3. Assessment comparability

As summarized in Table 4, positive assessments of NRT aligned with cieBOM's positive recommendations for the same treatment in two instances. In one case, the RCT-based and the NRT-based assessment were negative. In one further case, neither the RCT nor the matching NRT provided sufficient relevant data for a conclusive or preliminary assessment. Further, four positive recommendations based on the applicable RCT-PASKWIL criteria contrasted negative assessments for the matching respective NRTs. Four other cases saw the respective NRT-tested treatment being not assessable, while cieBOM gave a positive RCT-informed recommendation on the same treatment. No positive NRT-based assessments were found to contrast negative RCT-based recommendations.

4. Discussion

In our study, we matched NRTs on SACTs with succeeding RCTs underpinning cieBOM recommendations regarding treatment, treatment line, and patient population. Subsequently, we applied the NRT-PASKWIL criteria to the matching NRTs and conducted an assessment on the clinical benefit of the relevant SACTs based on these criteria. For 13 out of the 41 eligible RCTs, matching NRTs were found, out of which two (15%) were assessed positively, six (46%) negatively, and five (38%) were considered non-assessable based on the NRT-PASKWIL criteria. Two positive NRT-based assessments matched the corresponding positive cieBOM recommendations; in one case, both the assessment of an NRT outcome and the matching cieBOM recommendation was negative; and in one further case, the treatment could not be assessed based on the NRT data nor the RCT outcomes.

We found that the availability of matching NRTs was considerably limited. This may be explained by a potential tendency towards the exclusive publication of phase III results in journals compared to results from earlier trial phases. Moreover, SACT trials might undergo considerable changes between phases, such as eligibility criteria regarding indications and patient population becoming stricter, which might cause mismatches under the applied matching criteria.

Particularly for (novel) SACTs where RCT testing is unlikely due to the indication's rarity, a cumulative ORR/DoR criterion for assessing the treatment's clinical value could incentivize clinical researchers to publish study results where these endpoints are reported more consistently. This might facilitate a full assessment of the clinical benefit of the respective treatment. Further, clinical researchers might consider both endpoints when testing immunotherapies in particular, as they commonly have a lower ORR than targeted therapies, but a relatively long DoR [42]. However, assessing the clinical value of immunotherapies based on NRTs might have inherent limitations: negative outcomes of assessments regarding immunotherapies based on the ORR/DoR-related criteria (nivolumab as a second-/third-line treatment for accRCC [22]; pembrolizumab for advanced melanoma [18]) contrasted the NRT-based ESMO-MCBS scores. This means that assessing the clinical benefit of immunotherapies with ORR/DoR should be considered with caution. RCT-based assessments using PFS/OS and hazard ratios in relation to a comparator might provide a clearer picture. Our study exclusively included RCT results underpinning clinical value assessments by cieBOM. This limited the scope of comparing viable RCTs to matching NRTs. A broader scope (e.g., a longer time period for the inclusion of RCT-based cieBOM recommendations) might have increased our findings' robustness. Additionally, our study excluded results that were published as ESMO/ASCO abstracts only. Had such publications been included, there would have been three more NRTs matching the relevant RCTs. However, only one of these studies would have provided sufficient data on both ORR and DoR [43]. Therefore, the inclusion of those publications would likely not have significantly impacted the results of this study.

In practice, clinical value assessments by cieBOM based on NRTs are preliminary and subject to potential revision due to follow-up RCTs [44–46]. This approach reflects (early) marketing authorization decisions by EMA on SACTs based on uncertain evidence from single-arm and other NRTs: Conditional Marketing Authorization obligates manufacturers to provide additional, more robust data regarding efficacy, safety, and clinical endpoints for an SACT to receive full approval [47]. This indicates that NRT-based assessments of clinical value based on response-related criteria should in general be interpreted with caution. However, in rare cases with very promising NRT outcomes fulfilling the NRT-PASKWIL criteria, such as selpercatinib for RET-fusion-positive NSCLC, it was questioned whether performing an RCT was ethically justifiable [48,49]. Moreover, the results of our study suggest several practical implications related to the NRT-PASKWIL criteria and to assessing the clinical value of SACTs based on NRTs in general. First, NRT designs are generally fraught with multiple inherent uncertainty factors and shortcomings (e.g., small patient population, no comparator, shorter follow-up period). Combined with a considerable share of NRT results not being (fully) published, the viability of NRTs for clinical value assessments is substantially hampered. Secondly, the NRT-PASKWIL criteria are intended for treatments for rare indications, where RCTs are often not feasible due to a scarcity of patients. Overall, our findings suggest that the NRT-PASKWIL criteria are not suitable for NRT-based clinical value assessments of SACTs beyond that scope.

In conclusion, we showed that assessing the clinical benefit of SACTs tested in NRTs strongly depends on consistent and conclusive reporting of the relevant results. Our assessments based on the NRT-PASKWIL criteria did not consistently match the respective RCT-based cieBOM recommendations and yielded relatively more negative assessments. This may be explained by a disparity between the relevant appraisal criteria and inconsistent reporting in trial publications. NRT-based assessments are only preliminary and should not be regarded as definite. Moreover, our study suggests that the NRT-PASKWIL criteria are not suitable for clinical value assessments of SACTs beyond the scope of rare indications. However, the current course of action by EMA [4–7] may lead to a trend towards the NRT-PASKWIL criteria being applied more frequently. Therefore, staying mindful of the challenges inherent to NRT-based recommendations remains important.

Statement of ethics

No ethical approval was required for this study.

Funding

This study was funded by a grant from Stichting Kwaliteitsgelden Medisch Specialisten.

CRediT authorship contribution statement

N.S.H. Xander: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Formal analysis,

Conceptualization. **B. Leeneman**: Writing – review & editing, Methodology, Conceptualization. **A.M.C. Dingemans**: Writing – review & editing, Methodology, Conceptualization. **W.E. Fiets:** Writing – review & editing, Methodology, Conceptualization. **W.K. de Jong:** Writing – review & editing, Methodology, Conceptualization. **N.E.M. Uyl:** Writing – review & editing, Methodology, Conceptualization. **A.N.M. Wymenga:** Writing – review & editing, Methodology, Conceptualization. **A.K.L. Reyners:** Writing – review & editing, Methodology, Conceptualization. **C.A. Uyl-de Groot:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2024.114262.

References

[1] Farina A, Moro F, Fasslrinner F, Sedghi A, Bromley M, Siepmann T. Strength of clinical evidence leading to approval of novel cancer medicines in Europe: A systematic review and data synthesis. Pharm Res Perspect 2021;9(4):1–10. https://doi.org/10.1002/prp2.816.

[2] Spieth PM, Kubasch AS, Isabel Penzlin A, et al. Randomized clinical trials - a matter of design. Neuropsychiatr Dis Treat 2016;12:1341–9. https://doi.org/10.2147/NDT.S101938.

[3] Tannock IF, Amir E, Booth CM, et al. Relevance of randomised controlled trials in oncology. Lancet Oncol 2016;17(12):e560–7. https://doi.org/10.1016/S1470-2045(16)30572-1.

[4] Goring S, Taylor A, Müller K, et al. Characteristics of non-randomised studies using comparisons with external controls submitted for regulatory approval in the USA and Europe: a systematic review. BMJ Open 2019;9(2):1–11. https://doi.org/10.1136/bmjopen-2018-024895.

[5] Vreman RA, Bouvy JC, Bloem LT, et al. Weighing of evidence by health technology assessment bodies: retrospective study of reimbursement recommendations for conditionally approved drugs. Clin Pharm Ther 2019;105(3):684–91. https://doi.org/10.1002/cpt.1251.

[6] Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: Analysis of EMA and FDA approvals 1999-2014. BMJ Open 2016;6(6). https://doi.org/10.1136/bmjopen-2016-011666.

[7] Tenhunen O, Lasch F, Schiel A, Turpeinen M. Single-arm clinical trials as pivotal evidence for cancer drug approval: a retrospective cohort study of centralized European marketing authorizations between 2010 and 2019. Clin Pharm Ther 2020;108(3):653–60. https://doi.org/10.1002/cpt.1965.

[8] Nederlandse Vereniging voor Medische Oncologie (NVMO). Over de adviezen - NVMO. https://www.nvmo.org/over-de-adviezen/.

[9] Nederlandse Vereniging voor Medische Oncologie (NVMO). Over de commissie BOM. Nederlandse Vereniging voor Medische Oncologie [Internet]. ⟨https://www.nvmo.org/nvmo/commissie-bom/over-de-commissie-bom/⟩.

[10] Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European society for medical oncology magnitude of clinical benefit scale (ESMO-MCBS). Ann Oncol 2015;26(8):1547–73. https://doi.org/10.1093/annonc/mdv249.

[11] Cherny NI, Dafni U, Bogaerts J, et al. ESMO-magnitude of clinical benefit scale version 1.1. Ann Oncol 2017;28(10):2340–66. https://doi.org/10.1093/annonc/mdx310.

[12] Leeneman B, Xander NSH, Fiets WE, et al. Assessing the clinical benefit of systemic anti-cancer treatments in the Netherlands: the impact of different thresholds for effectiveness. Eur J Cancer 2024;202:114002. https://doi.org/10.1016/j.ejca.2024.114002.

[13] Danen A. 'We hebben de lat hoog gelegd.' Medische Oncologie. doi:10.24078/onco.2021.5.127336.

[14] NVMO-commissie B.O.M. BOM Archieven - NVMO. ⟨https://www.nvmo.org/bom-type/bom/?order=disease⟩.

[15] Kurzrock R, Sherman SI, Ball DW, et al. Activity of XL184 (cabozantinib), an oral tyrosine kinase inhibitor, in patients with medullary thyroid cancer. J Clin Oncol 2011;29(19):2660–6. https://doi.org/10.1200/JCO.2010.32.4145.

[16] McGonigle KF, Muntz HG, Vuky J, et al. Combined weekly topotecan and biweekly bevacizumab in women with platinum-resistant ovarian, peritoneal, or fallopian tube cancer: results of a phase 2 study. Cancer 2011;117(16):3731–40. https://doi.org/10.1002/cncr.25967.

[17] Monk BJ, Sill MW, Burger RA, Gray HJ, Buekers TE, Roman LD. Phase II trial of bevacizumab in the treatment of persistent or recurrent squamous cell carcinoma of the cervix: A gynecologic oncology group study. J Clin Oncol 2009;27(7):1069–74. https://doi.org/10.1200/JCO.2008.18.9043.

[18] Hamid O, Robert C, Daud A, et al. Safety and tumor responses with lambrolizumab (Anti–PD-1) in melanoma. N Engl J Med 2013;369(2):134–44. https://doi.org/10.1056/nejmoa1305133.

[19] Ribas A, Gonzalez R, Pavlick A, et al. Combination of vemurafenib and cobimetinib in patients with advanced BRAFV600-mutated melanoma: a phase 1b study. Lancet Oncol 2014;15(9):954–65. https://doi.org/10.1016/S1470-2045(14)70301-8.

[20] Cabanillas ME, Schlumberger M, Jarzab B, et al. A phase 2 trial of lenvatinib (E7080) in advanced, progressive, radioiodine-refractory, differentiated t. Cancer 2015;121(16):2749–56. https://doi.org/10.1002/cncr.29395.

[21] Reck M, Kaiser R, Eschbach C, et al. A phase II double-blind study to investigate efficacy and safety of two doses of the triple angiokinase inhibitor BIBF 1120 in patients with relapsed advanced non-small-cell lung cancer. Ann Oncol 2011;22(6):1374–81. https://doi.org/10.1093/annonc/mdq618.

[22] Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for metastatic renal cell carcinoma: results of a randomized phase II trial. J Clin Oncol 2015;33(13):1430–7. https://doi.org/10.1200/JCO.2014.59.0703.

[23] Choueiri TK, Pal SK, McDermott DF, et al. A phase I study of cabozantinib (XL184) in patients with renal cell cancer. Ann Oncol 2014;25(8):1603–8. https://doi.org/10.1093/annonc/mdu184.

[24] Vahdat LT, Pruitt B, Fabian CJ, et al. Phase II study of eribulin mesylate, a halichondrin B analog, in patients with metastatic breast cancer previously treated with an anthracycline and a taxane. J Clin Oncol 2009;27(18):2954–61. https://doi.org/10.1200/JCO.2008.17.7618.

[25] Senzer NN, Kaufman HL, Amatruda T, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. J Clin Oncol 2009;27(34):5763–71. https://doi.org/10.1200/JCO.2009.24.3675.

[26] Goss G, Tsai CM, Shepherd FA, et al. Osimertinib for pretreated EGFR Thr790Met-positive advanced non-small-cell lung cancer (AURA2): a multicentre, open-label, single-arm, phase 2 study. Lancet Oncol 2016;17(12):1643–52. https://doi.org/10.1016/S1470-2045(16)30508-3.

[27] Bruix J, Tak WY, Gasbarrini A, et al. Regorafenib as second-line therapy for intermediate or advanced hepatocellular carcinoma: Multicentre, open-label, phase II safety study. Eur J Cancer 2013;49(16):3412–9. https://doi.org/10.1016/j.ejca.2013.05.028.

[28] Elisei R, Schlumberger MJ, Müller SP, et al. Cabozantinib in progressive medullary thyroid cancer. J Clin Oncol 2013;31(29):3639–46. https://doi.org/10.1200/JCO.2012.48.4659.

[29] Pujade-Lauraine E, Hilpert F, Weber B, et al. Bevacizumab combined with chemotherapy for platinum-resistant recurrent ovarian cancer: the AURELIA open-label randomized phase III trial. J Clin Oncol 2014;32(13):1302–8. https://doi.org/10.1200/JCO.2013.51.4489.

[30] Tewari KS, Sill MW, Long HJ, et al. Improved survival with bevacizumab in advanced cervical cancer. N Engl J Med 2014;370(8):734–43. https://doi.org/10.1056/nejmoa1309748.

[31] Robert C, Schachter J, Long GV, et al. Pembrolizumab versus Ipilimumab in advanced melanoma. N Engl J Med 2015;372(26):2521–32. https://doi.org/10.1056/nejmoa1503093.

[32] Larkin J, Ascierto PA, Dréno B, et al. Combined vemurafenib and cobimetinib in BRAF -mutated melanoma. N Engl J Med 2014;371(20):1867–76. https://doi.org/10.1056/nejmoa1408868.

[33] Schlumberger M, Tahara M, Wirth LJ, et al. Lenvatinib versus placebo in radioiodine-refractory thyroid cancer. N Engl J Med 2015;372(7):621–30. https://doi.org/10.1056/nejmoa1406470.

[34] Reck M, Kaiser R, Mellemgaard A, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): A phase 3, double-blind, randomised controlled trial. Lancet Oncol 2014;15(2):143–55. https://doi.org/10.1016/S1470-2045(13)70586-2.

[35] Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 2015;373(19):1803–13. https://doi.org/10.1056/nejmoa1510665.

[36] Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus everolimus in advanced renal-cell carcinoma. N Engl J Med 2015;373(19):1814–23. https://doi.org/10.1056/nejmoa1510016.

[37] Choueiri TK, Escudier B, Powles T, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. Lancet Oncol 2016;17(7):917–27. https://doi.org/10.1016/S1470-2045(16)30107-3.

[38] Kaufman PA, Awada A, Twelves C, et al. Phase III open-label randomized study of eribulin mesylate versus capecitabine in patients with locally advanced or metastatic breast cancer previously treated with an anthracycline and a taxane. J Clin Oncol 2015;33(6):594–601. https://doi.org/10.1200/JCO.2013.52.4892.

[39] Andtbacka RHI, Ross M, Puzanov I, et al. Patterns of clinical response with talimogene laherparepvec (T-VEC) in patients with melanoma treated in the OPTiM Phase III clinical trial. Ann Surg Oncol 2016;23(13):4169–77. https://doi.org/10.1245/s10434-016-5286-0.

[40] Mok TS, Wu YL, Ahn MJ, et al. Osimertinib or platinum–pemetrexed in EGFR T790M–positive lung cancer. N Engl J Med 2017;376(7):629–40. https://doi.org/10.1056/nejmoa1612674.

[41] Bruix J, Qin S, Merle P, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised,

double-blind, placebo-controlled, phase 3 trial. Lancet 2017;389(10064):56–66. https://doi.org/10.1016/S0140-6736(16)32453-9.
[42] Wargo JA, Cooper ZA, Flaherty KT. Universes collide: combining immunotherapy with targeted therapy for cancer. Cancer Discov 2014;4(12):1377–86. https://doi.org/10.1158/2159-8290.CD-14-0477.
[43] Kim DW, Ahn MJ, Shi Y, et al. Results of a global Phase II study with crizotinib in advanced ALK-positive non-small-cell lung cancer (NSCLC). Ann Oncol 2012;23:xi32–3. https://doi.org/10.1016/S0923-7534(20)32006-8.
[44] NVMO-commissie BOM. Selpercatinib als tweedelijnsbehandeling bij gemetastaseerd RET-fusiepositief niet-kleincellig longcarcinoom. Med Oncol 2022;25(1):35–8.
[45] NVMO-commissie BOM. Entrectinib bij het niet-kleincellig longcarcinoom met een ROS1-genfusie. Med Oncol 2022;25(1):23–7.
[46] NVMO-commissie BOM. Entrectinib bij solide tumoren met een NTRK-genfusie. Med Oncol 2022;25(1):17–21.
[47] Michaeli DT, Michaeli T, Albers S, Boch T, Michaeli JC. Special FDA designations for drug development: orphan, fast track, accelerated approval, priority review, and breakthrough therapy. Published online November 14 Eur J Health Econ 2023. https://doi.org/10.1007/s10198-023-01639-x. Published online November 14.
[48] Zhou C, Solomon B, Loong HH, et al. First-line selpercatinib or chemotherapy and pembrolizumab in *RET* fusion–positive NSCLC. N Engl J Med 2023;389(20):1839–50. https://doi.org/10.1056/NEJMoa2309457.
[49] Dingemans AMC, Smit EF, Hendriks LEL. Selpercatinib or chemotherapy in *RET* FUsion–positive NSCLC. N Engl J Med 2024;390(4):381–2. https://doi.org/10.1056/NEJMc2314327.