PharmacoEconomics
https://doi.org/10.1007/s40273-024-01447-y

SYSTEMATIC REVIEW

A Systematic Literature Review of Modelling Approaches to Evaluate the Cost Effectiveness of PET/CT for Therapy Response Monitoring in Oncology

Sietse van Mossel[1,2] · Rafael Emilio de Feria Cardet[3] · Lioe-Fee de Geus-Oei[1,2,4] · Dennis Vriens[5] · Hendrik Koffijberg[6] · Sopany Saing[3,6]

Accepted: 9 October 2024

Abstract

Background and Objective This systematic literature review addresses model-based cost-effectiveness studies for therapy response monitoring with positron emission tomography (PET) generally combined with low-dose computed tomography (CT) for various cancer types. Given the known heterogeneity in therapy response events, studies should consider patient-level modelling rather than cohort-based modelling because of its flexibility in handling these events and the time to events. This review aims to identify the modelling methods used and includes a systematic assessment of the assumptions made in the current literature.

Methods This study was conducted and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement. Information sources included electronic bibliographic databases, reference lists of review articles and contact with experts in the fields of nuclear medicine, health technology assessment and health economics. Eligibility criteria included peer-reviewed scientific publications and published grey literature. Literature searches, screening and critical appraisal were conducted by two reviewers independently. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) were used to assess the methodological quality. The Bias in Economic Evaluation (ECOBIAS) checklist was used to determine the risk of bias in the included publications.

Results The search results included 2959 publications. The number of publications included for data extraction and synthesis was ten, representing eight unique studies. These studies addressed patients with lymphoma, advanced head and neck cancers, brain tumours, non-small cell lung cancer and cervical cancer. All studies addressed response to chemotherapy. No study evaluated response to immunotherapy. Most studies positioned PET/CT as an add-on modality and one study positioned PET/CT as a replacement for conventional imaging (X-ray and contrast-enhanced CT). Three studies reported decision-tree structures, four studies reported cohort-level state-transition models and one study reported a partitioned survival model. No patient-level models were reported. The simulation horizons adopted ranged from 1 year to lifetime. Most studies reported a probabilistic analysis, whereas two studies reported a deterministic analysis only. Two studies conducted a value of information analysis. Multiple studies did not adequately discuss model-specific aspects of bias. Most importantly and regularly observed were a high risk of structural assumptions bias, limited simulation horizon bias and wrong model bias.

Conclusions Model-based cost-effectiveness analysis for therapy response monitoring with PET/CT was based on cohorts of patients instead of individual patients in the current literature. Therefore, the heterogeneity in therapy response events was commonly not addressed appropriately. Further research should include more advanced and patient-level modelling approaches to accurately represent the complex context of clinical practice and, therefore, to be meaningful to support decision making.

Registration This review is registered in PROSPERO, the international prospective register of systematic reviews funded by the National Institute for Health Research, with CRD42023402581.

Extended author information available on the last page of the article

Published online: 03 November 2024

Key Points for Decision Makers

There have only been a few published health economic evaluations of the use of positron emission tomography/computed tomography in therapy response monitoring, since its introduction for oncology in 2001.

Simulation models evaluating the cost effectiveness of therapy response monitoring with positron emission tomography/computed tomography are based on cohorts of patients rather than individual patients. The heterogeneity in therapy response events is, therefore, commonly not appropriately addressed.

Future research should include more advanced and/or patient-level modelling approaches to accurately reflect therapy response monitoring in clinical oncology practice and, therefore, to be meaningful to support decision making.

1 Introduction

The past decade has seen compelling developments in systemic cancer therapies resulting in improved patient outcomes. At the same time, new challenges emerged including the need for the oncological community to reconsider the conventional ways of therapy response monitoring. Measurement of response to therapy—in oncology often pursued by serial transactional imaging using computed tomography (CT), magnetic resonance imaging (MRI) or positron emission tomography (PET) in combination with low-dose CT and/or serial follow-up of a serological biomarkers—is used both in studies and in clinical routine as an immediate marker of patient prognosis (e.g. disease-free or overall survival [OS]), and provides the opportunity to change clinical decision making [1, 2]. Historically, based on morphological imaging (CT or MRI), response to therapy is defined by changes in the tumour diameter and number of lesions. A shrinkage in tumour size above a preset threshold and/or a decrease in the number of lesions categorises patients as being responsive to therapy [3, 4]. The opposite, an increased diameter above a preset threshold and/or increase in the number of lesions, signifies progressive disease. However, flare reactions or pseudoprogression (i.e. an initial increase in tumour size preceding response) and dissociated or mixed responses are also regularly observed, mainly in immunotherapeutics [5–7]. Therefore, morphological imaging (CT or MRI) might not provide the best immediate marker of the actual response to therapy and, thus, prognosis. Consequently, correct continuation of effective therapy, but also timely discontinuation of ineffective therapy and shifting patients to alternative therapy options, is suboptimal in a meaningful subgroup of patients.

Response to therapy is determined with the Response Evaluation Criteria in Solid Tumours (RECIST) and adaptions to RECIST have been made for immunomodulatory drugs in 2009 (iRECIST) [3]. The RECIST Working Group 2009 also included an examination on whether it was convenient to move from morphological (CT or MRI) to functional imaging using PET (i.e. molecular imaging considering a tumour's metabolism such as glucose uptake and protein expression). In 2009, there was insufficient standardisation and evidence available to abandon a morphological assessment. By now, several qualitative and quantitative criteria for response assessment with PET are thoroughly validated and a metabolic response (shutdown of metabolism) generally precedes a morphological response (tumour shrinkage) [5–9]. Therefore, PET biomarkers indicating a metabolic response—generally combined with low-dose CT to preserve detailed information on location, size and tissue characteristics [10]—improve clinical decision making to continue, timely change or discontinue therapy.

The high costs of therapy and the health burden to patients emphasise the added value of more accurate and early response monitoring with PET/CT. Depending on the positioning of PET/CT in the care pathway, it can be considered as a replacement scan for purely morphological imaging (CT or MRI) or as an early interim or add-on scanner option [11]. When morphological imaging accuracy is unsatisfactory, PET/CT should be considered as a replacement modality. Early PET/CT is relevant to support timely shifting or discontinuing therapy by shortening the lead time to response assessment. The considerable cost of PET/CT itself suggests that insight into the cost effectiveness of PET/CT is needed to inform optimal usage [2, 12]. Randomised clinical trials (RCTs) may support a cost-effectiveness analysis. However, RCTs are not always recommended because they are costly and time consuming. They are also considered less feasible given the large number of available response assessment strategies that can be compared, and the limited time horizon of an RCT may not allow the evaluation of long-term costs and health outcomes. [13–15] These limitations are apparent from the low number of RCTs for therapy response monitoring in the current literature [2].

As an alternative for conducting RCTs, simulation modelling can be applied to estimate the cost and health benefits of using PET/CT for therapy response monitoring [14–16]. Consensus guidelines provide direction to select an appropriate model structure given the decision problem and available evidence [17–19]. Different modelling approaches can be applied including state-transition models (STM), partitioned

survival models (PSM) and discrete event simulations. There is sufficient awareness among policy makers and decision makers that the choice of model structure and evidence used to parametrise the model may affect cost-effectiveness outcome estimates [20, 21]. Given the known heterogeneity in therapy response events, it is suggested that modelling studies consider patient-level modelling rather than cohort-based modelling because of its flexibility in handling these events and the time to events [22]. Accordingly, this study systematically assesses whether published simulation models for a cost-effectiveness analysis adequately reflect heterogeneity in events, and whether patient-level models are available in the current literature.

2 Methods

This systematic literature review is conducted and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [23]. This review is also registered in PROSPERO, the international prospective register of systematic reviews funded by the National Institute for Health Research, with CRD42023402581.

2.1 Eligibility Criteria

Eligibility criteria covered peer-reviewed scientific publications and published grey literature. Eligible scientific publications were original research articles. Excluded scientific publications were review articles, editorial letters, commentaries, conference reports, consensus and best-practice guidelines. Only studies published from the year 2001 onwards were included. This time horizon was selected because the year 2001 marks the introduction of integrated PET/CT for clinical oncology [24]. The focus was on studies evaluating PET-based tumour response-adaptive systemic therapy approaches. Hence, studies were excluded if the role of PET/CT in cancer screening, pre-treatment diagnosis and staging, or re-staging after radical treatment was considered. In addition, eligible studies should have reported on full health economic evaluations defined as cost-effectiveness, cost-utility and cost-benefit studies. Therefore, costing studies and budget impact analyses were excluded. Additionally, studies reporting trial-based economic evaluations were excluded.

2.2 Information Sources and Search Strategy

Information sources included electronic bibliographic databases, reference lists of review articles, and contact with study investigators and experts in nuclear medicine, health technology assessment (HTA) and health economics. The databases and search strategies were discussed with an information specialist and exploratively tested to verify that all relevant papers already known to the study investigators were retrieved. At the least, relevant search results in the disease areas of malignant lymphoma and advanced head and neck cancers were expected. A comprehensive search in four bibliographic databases was conducted including Scopus and PubMed as well as the HTA domain-specific international HTA database and the National Health Service Economic Evaluation Database. Search dates were from 16 February, 2023 to 24 February, 2023. Searches were updated to identify and retrieve any further studies for inclusion before the full-text review (3 May, 2023), data extraction (17 July, 2023), and quality and risk of bias assessments (26 January, 2024). The databases were searched using free-text terms and medical subject headings (MeSH terms). Free-text terms included cost-effectiveness and cost-utility, health economic evaluation and HTA, PET/CT, cancer or malignancy, and systemic therapies including immunotherapy, chemotherapy and hormonotherapy. The search queries are reported in the ESM. Exclusion criteria regarding the availability of full texts and non-English language were not enforced during the searches to avoid erroneous exclusion of eligible publications. A machine translation engine that leverages convolutional neural networks to convert text (DeepL Translator) was found to be a viable tool for translating database searches into English to facilitate the study selection process [25].

2.3 Selection Process

Following the PRISMA guidelines, literature searches, screening and critical appraisal were conducted by multiple reviewers. Covidence systematic review software and automation tools were used for the screening, full-text review and data extraction process [26]. Automation was used for de-duplication of references, for highlighting inclusion and exclusion keywords to assist with the screening, and for adding study tags to assist with the filtering of reasons for final inclusion or exclusion of publications.

Reviewers SvM and RFC independently screened titles and abstracts from all searches and reviewed the full texts of the selected publications for final inclusion or exclusion. Disagreements between the reviewers during both the initial screening and full-text review were resolved by consensus after consulting SS. Studies excluded during the full-text review were categorised according to the following categories: non-English articles, ineligible article type, no comparison of diagnostics, PET/CT used in non-oncology care, no assessment of mass forming tumours, no assessment of response to therapy, no simulation modelling applied and no full health economic evaluation performed. Articles discussing malignant lymphoma might be included as

hematogenous tumours may form solid masses. Additionally, it is worth noting that articles excluded at the stage of study selection because of language will be included in the list of excluded studies. This allows transparency regarding the number of eligible reports in other languages and allows future review authors to investigate these studies when checking prior literature reviews on the same topic [27].

2.4 Data Extraction and Synthesis

Reviewer SvM performed the full data extraction from all included publications. General information was extracted for each of the following aspects: author list, year of publication, journal of publication, location of publication, funding sources, the aim of the study, study population, imaging used, tumour biomarkers and radiopharmaceuticals, diagnostic timing and time interval including recurrent cycles, alternative diagnostics, diagnostic performance, consequences of incorrect classifications, therapy strategies, therapy line of focus, type of health economic evaluation, perspective of analysis (hospital, insurance company or societal), discounting, estimated costs and health effects.

Furthermore, simulation modelling information was extracted for each of the following aspects: modelling technique, model structure, level of modelling (cohort or patient-based), evidence sources, comparator strategy, reflection of heterogeneity, extrapolation methods, competing risks, time-cycle length, or time-to-event distributions and simulation time horizon. Additionally, the approaches used to characterise uncertainty were extracted. The consideration of a subgroup analysis, scenario analysis, probabilistic analysis, value of information (VOI) analysis and main areas of structural uncertainty including major assumptions and limitations mentioned was assessed. In addition, the reporting of validation was evaluated using the Assessment of the Validation Status of Health-Economic Decision Models (AdViSHE) consensus checklist [28].

2.5 Quality Assessment

Reviewers SvM and RFC independently performed a structured quality assessment of the included publications. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) were used to assess the methodological quality of the included publications [29]. The assessment criteria for the individual items from the CHEERS checklist were determined before having executed the database searches and was also reported in the PROSPERO registration (CRD42023402581). More specifically, the title and abstract were correctly reported if they defined all imaging, tumour biomarkers/tracers and the therapy line of focus (abstract only), the 'selection of health outcomes' item was correctly reported if the choice of primary health outcomes was explained where this was not quality-adjusted life-years, the 'measurement of effectiveness' item was correctly reported if a search and analysis of relevant studies was described, and uncertainty in the results was considered to be correctly reported if confidence intervals were provided for the primary health economic outcomes or if an incremental cost-effectiveness plane was presented including the distribution of outcomes or a confidence ellipse. Furthermore, items regarding the inclusion of patient preferences and the reflection of heterogeneity in the outcomes could be not applicable. The template is provided in the ESM.

2.6 Risk of Bias Assessment

Reviewers SvM and RFC independently performed a structured risk of bias assessment of the included publications. The Bias in Economic Evaluation (ECOBIAS) checklist was used to assess the risk of bias in the included publications [30]. The checklist contains 22 criteria and includes 11 criteria related to model-specific aspects of biases. The assessment criteria for the individual items from the ECOBIAS checklist were determined before having executed the database searches and was also reported in the PROSPERO registration (CRD42023402581). Specifically, model-specific biases were classified as biases related to model structure, data sources or internal consistency. Biases related to the model structure included structural assumptions bias, comparator bias, wrong model bias and limited simulation horizon bias. Biases related to data sources included biases related to health effects, cost data and limited scope bias. The template is provided in the ESM.

3 Results

3.1 Search Results

The PRISMA flowchart shows the search results (Fig. 1). A total of 2959 publications were obtained from the bibliographic database searches. There were 1914 records identified from the Scopus database, 901 records from the PubMed database, 30 records from the international HTA database and 114 records from the National Health Service Economic Evaluation Database. A total of 561 duplicate records were removed before screening. After de-duplication, 2398 publications were included in the title and abstract screening. A total of 2341 publications were excluded during the title and abstract screening and 57 publications were included in the full-text review. The three main reasons for exclusion were the absence of a full health economic evaluation, the absence of simulation modelling and the absence of therapy response monitoring. Based on a full-text review, a total of 47 further publications were excluded. Reasons for

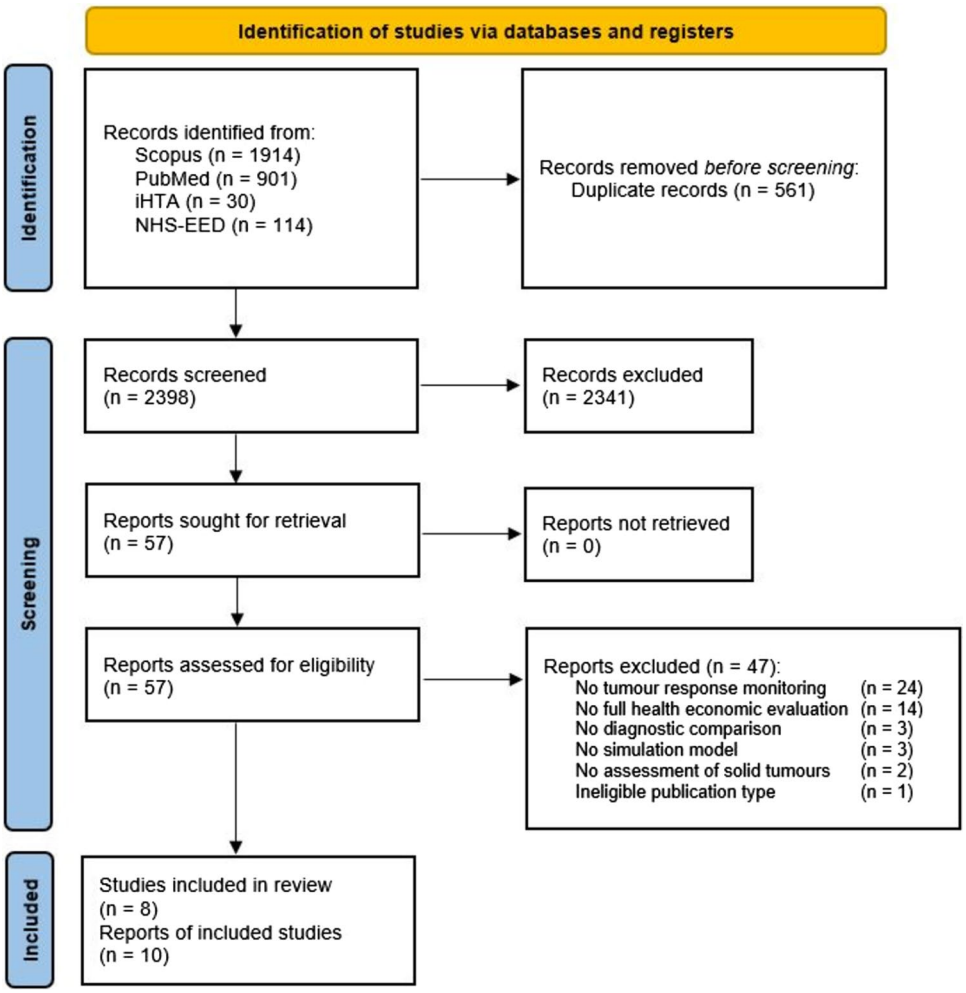

Fig. 1 Preferred reporting items for systematic reviews and meta-analyses (PRISMA) 2020 statement flowchart visualises the search results [23]. A total of eight unique studies, for which ten reports were consulted, were included for review. Note that reasons for exclusion during eligibility assessment are also included. *iHTA* international health technology assessment, *NHS-EED* National Health Service Economic Evaluation Database

exclusion are reported in Fig. 1. The number of publications included for data extraction and synthesis was ten [31–40]. Two studies were published in a scientific journal and as a full report within the UK National Institute for Health and Care Research HTA programme [31, 34, 39, 40]. For these studies, data were extracted based on the information available in all relevant reports, resulting in a final sample of eight unique studies [32–39]. For two studies, data were extracted only for the subgroup of patients with advanced disease who received systemic therapy [34, 37].

3.2 Study Characteristics

Most studies were conducted in Europe [33–35, 37–39], one study in the USA [32], and one study in Canada [36]. All studies were published in a clinical journal. Most studies adopted a healthcare perspective [32, 34, 36–39], while only two Dutch studies considered a societal perspective covering potential productivity losses [33, 35]. One study studied patients with malignant lymphoma [33], four studies studied patients with advanced head and neck cancers [32, 35, 36, 39], one study studied patients with brain tumours [38], one study studied patients with non-small cell lung cancer [37], and one study studied patients with late-stage and persistent cervical cancer [34]. All studies studied responses to chemotherapy. Prescribed medication differed between the studies and was not always reported. Most studies considered first-line systemic therapy [32–37, 39], but also second-line therapy [37] and adjuvant therapy [38] were considered. The simulation time horizons adopted ranged from 1 year to lifetime. Three studies considered a lifetime horizon [34, 36, 39], two studies considered a fixed horizon of 5 years [33, 37], and one study considered a fixed horizon of 1 year [32]. The remaining two studies considered a horizon shorter than 1 year and adaptive to the therapy duration [35, 38]. A study-level overview is provided in Table 1.

3.3 Imaging Characteristics

All studies compared PET/CT using [^{18}F]Fluorodeoxyglucose (FDG) as PET-radiopharmaceutical with morphological imaging [33–37, 39], except for one study that

Table 1 Study-level overview of general study characteristics including healthcare perspective, therapy information and simulation time horizon

Study	Location	Study objective	Patient population	Therapy strategy	Therapy line	Prescribed medication	Comparator imaging	PET radiopharmaceutical	Histologically proven cancer?	Time to imaging	Time horizon	Funding
Van Loon (2010)[a] [37]	NL	To evaluate the cost effectiveness of a PET/CT-based follow-up as opposed to a conventional follow-up in patients with NSCLC treated with CRT	Patients with advanced staged NSCLC	CRT	[1] First-line, [2] second-line	[1] Docetaxel, [2] erlotinib	[1] X-ray, [2] CT, [3] PET/CT	[18F]FDG	Not reported	End-of-therapy imaging[d]	5 years	Not reported
Rabalais (2012) [32]	US	To study the cost effectiveness of PET/CT for the management of the neck in patients with N2 neck cancer treated with CRT	Patients with advanced nodal neck cancer	CRT	First-line	Not reported	[1] None, [2] CT, [3] PET/CT	Not reported	Not reported	End-of-therapy imaging[d]	1 year	Non-industry
Auguste & Meads (2014)[b] [34]	UK	To undertake a cost-effectiveness analysis that compares PET/CT plus standard practice with standard practice alone in the diagnosis of recurrent or persistent cervical cancer during routine surveillance and a follow-up in women treated with CRT	Patients with persistent and late-stage cervical cancer	CRT	First-line	Carboplatin/5-fluorouracil	[1] MRI and/or CT, [2] MRI and/or CT plus PET/CT	[18F]FDG	Yes	End-of-therapy imaging[d]	Lifetime	Non-industry

Table 1 (continued)

Study	Location	Study objective	Patient population	Therapy strategy	Therapy line	Prescribed medication	Comparator imaging	PET radiopharmaceutical	Histologically proven cancer?	Time to imaging	Time horizon	Funding
Greuter (2017) [35]	NL	To assess the effects and costs of four response evaluation strategies including PET/CT to detect local residual disease in patients with advanced squamous cell OPC treated with CRT	Patients with advanced staged squamous cell OPC	CRT	First-line	Not reported	[1] None, [2] MRI, [3] PET/CT	[18F]FDG	Yes	End-of-therapy imaging[d]	Therapy duration	Non-industry
Smith & Mehanna (2017) [39]	UK	To provide vital evidence for the cost effectiveness of a PET/CT guided management strategy for patients with N2 or N3 head and neck cancers treated with CRT	Patients with advanced nodal head and neck cancers	CRT	First-line	TPF	[1] None, [2] PET/CT	[18F]FDG	Yes	End-of-therapy imaging[d]	Lifetime	Non-industry
Fu (2021) [36]	CAN	To compare the long-term cost utility of CT surveillance, PET/CT–guided surveillance and PRND for the management of patients with advanced nodal HPV+ OPC treated with CRT	Patients with HPV+ advanced nodal OPC	CRT	First-line	Not reported	[1] None, [2] CT, [3] PET/CT	[18F]FDG	Not reported	End-of-therapy imaging[d]	Lifetime	Non-industry

Table 1 (continued)

Study	Location	Study objective	Patient population	Therapy strategy	Therapy line	Prescribed medication	Comparator imaging	PET radiopharmaceutical	Histologically proven cancer?	Time to imaging	Time horizon	Funding
Greuter (2022) [33]	NL	To assess the cost effectiveness of shortening therapy duration based on interim PET/CT scans in patients with DLBCL treated with R-CHOP	Patients with DLBCL	CHEMO	First-line	R-CHOP	[1] PET/CT, [2] PET/CT plus interim PET/CT	[18F]FDG	Not reported	Early imaging after two cycles of therapy	5 years	Non-industry[c]
Rosen (2022) [38]	GER	To determine the cost effectiveness of FET PET compared with conventional MRI for early identification of responders to adjuvant CHEMO in patients with glioma	Patients with glioblastoma	CRT	Adjuvant	Temozolomide	[1] Interim MRI, [2] interim PET	[18F]FET	Yes	Early imaging after two cycles of therapy	Therapy duration	Not reported

CAN Canada, *CHEMO* chemotherapy, *CT* computed tomography, *CRT* chemoradiotherapy, *DLBCL* diffuse large B-cell lymphoma, *FDG* fluorodeoxyglucose, *FET* fluoroethyl-L-tyrosine, *GER* Germany, *HPV+* human papillomavirus-positive, *MRI* magnetic resonance imaging, *NL* Netherlands, *NSCLC* non-small cell lung cancer, *OPC* oropharyngeal cancer, *PET* positron emission tomography, *R-CHOP* chemotherapy regimen of rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone, *TPF* chemotherapy regimen consisting of docetaxel, platinum and 5-fluorouracil, *UK* United Kingdom, *US* United States

[a]For the study of Van Loon et al., data were only extracted for the subgroup of patients who received chemoradiotherapy

[b]For the study of Auguste and Meads et al., data were only extracted for the subgroup of patients with late-stage and persistent disease who received chemoradiotherapy

[c]The PETAL trial was not funded from industry but multiple authors from the author list and part of the PETRA consortium disclosed to receive funding for consortium-related research activities

[d]End-of-therapy imaging was always performed for therapy response monitoring and conducted within 3 months after stopping therapy

considered stand-alone PET using [^{18}F]Fluoroethyl-L-tyrosine as a PET radiopharmaceutical [38]. One study did not report the PET radiopharmaceutical used [32]. In six studies, therapy response monitoring was scheduled within 3 months after the end of treatment [32, 34–37, 39]. In two studies, early therapy response monitoring was scheduled after two initial cycles of therapy [33, 38]. The performance of PET/CT was compared with contrast-enhanced CT [32, 34, 36, 37] or MRI [34, 35, 38]. One study included X-ray imaging as a relevant comparator [37]. Imaging test results were supported with histological proof in four studies [34, 35, 38, 39]. Histological proof was not reported in the other studies [32, 33, 36, 37]. One study positioned PET/CT as a replacement scan for morphological imaging [37], while the other studies positioned PET/CT as add-on scans [32–36, 38, 39]. As PET/CT is an expensive imaging modality, hospitals want to restrict its use as an add-on scan to patients who likely benefit from the imaging result. In general, add-on scans increase the sensitivity of an existing care pathway but potentially at the cost of specificity [11]. Depending on the positioning of PET/CT within the care pathway, diagnostic performance measures of sensitivity, specificity, positive predictive value or negative predictive value were alternately used in the simulation models.

Imaging test results impact therapy selection and further patient management. Four studies reported imaging-based stratification of patients to subsequent therapies [34, 36, 37, 39]. These studies distinguished curatively and palliatively intended therapy options and best supportive care. Moreover, four studies reported the avoidance of overtreatment [32, 33, 35, 38]. In these studies, patients were prevented from receiving futile or unnecessary therapy. Furthermore, five studies estimated patients' survival after sequential imaging and therapy [33, 34, 36, 37, 39]. Three studies reported 5-year OS rates [33, 34, 37], one study reported 3-year OS rates [34], and three studies reported 2-year OS rates [34, 36, 39]. Noteworthy, costing consequences of sequential imaging and therapy were commonly included as the lump sum cost per patient. A study-level overview of the reported imaging test results, treatment consequences and corresponding survival estimation is provided in Table 2.

3.4 Model Characteristics

Three studies reported decision-tree structures [32, 35, 38], four studies reported cohort-level STM [34, 36, 37, 39], and one study reported a traditional three-health-state PSM with progression-free, progression and death health states [33]. No patient-level STM, discrete event simulations or agent-based models were reported, which was unexpected given the heterogeneity observed in many of the considered patient populations. Of the five models

Table 2 Study-level overview of the reported imaging test results, therapy consequences and survival estimations

Study (year)	Role of PET/CT	Imaging statistics	Imaging test results	Consequences (therapy selection)	Consequences (survival estimation)
Van Loon (2010) [37]	Replacement scan	Not reported	[1] No progression, [2] symptomatic progression, [3] asymptomatic progression	[1] Curative intent, [2] palliative intent, [3] supportive care	5-Year OS
Rabalais (2012) [32]	Add-on scan	Not reported	[1] Response, [2] no response	Avoid neck dissection	N/A
Auguste and Meads (2014) [34]	Add-on scan	[1] Sens, [2] spec	[1] No progression, [2] symptomatic progression, [3] asymptomatic progression	[1] Curative intent, [2] palliative intent	[1] 2-year OS, [2] 3-year OS, [3] 5-year OS
Greuter (2017) [35]	Add-on scan	[1] Sens, [2] spec, [3] PPV, [4] NPV	[1] Response, [2] no response	Avoid examination under anaesthesia	N/A
Smith and Mehanna (2017) [39]	Add-on scan	[1] Sens, [2] spec, [3] PPV, [4] NPV	[1] Response, [2] no response	[1] Curative intent, [2] palliative intent	2-Year OS
Fu (2021) [36]	Add-on scan	Not reported	[1] Response, [2] no response	[1] Surgery, [2] salvage therapy, [3] palliative therapy	2-Year OS
Greuter (2022) [33]	Add-on scan	Not reported	[1] Early progression, [2] no early progression	[1] Stop therapy, [2] continue therapy	5-Year OS
Rosen (2022) [38]	Add-on scan	Not reported	[1] Early response, [2] no early response	[1] Stop therapy, [2] continue therapy	N/A

CT computed tomography, *N/A* not applicable, *NPV* negative predictive value, *OS* overall survival, *PET* positron emission tomography, *PPV* positive predictive value, *Sens* sensitivity, *Spec* specificity

explicitly considering time, one study considered a fixed cycle time of 6 months [37], two studies considered a fixed cycle time of 3 months [34, 36], one study considered a fixed cycle time of 1 month [39] and one study considered a fixed cycle time of 3 weeks [33]. Three studies used aggregated evidence from the scientific literature to populate their models [32, 34, 38], while five studies primarily used individual patient data (IPD) gathered from prospective studies, RCTs or national registries [33, 35–37, 39]. All studies provided a clear rationale for the selection of their evidence sources. Noteworthy, only four studies reported the software used for model building, parametrisation and simulation [34, 36, 38, 39]. A study-level overview is provided in Table 3. An overview of the conducted model validation is available in the Electronic Supplementary Material (ESM).

3.5 Health Economic Analysis

A cost-utility analysis assessing the cost per quality-adjusted life-year was performed in five studies [33, 34, 36, 37, 39]. To measure quality of life, two studies reported the use of the EuroQol 5D questionnaire [34, 39] and one study reported the use of the EORTC QLQ-C30 questionnaire [33]. The remaining two studies did not specify the questionnaire used [36, 37]. All cost-utility analyses were based on cohort-level STM or PSM. The willingness-to-pay (WTP) thresholds ranged from €20,000 to €80,000 per quality-adjusted life-year depending on the severity of disease. The discount rates ranged from 1.5 to 4%. Three cost-utility studies concluded that the use of PET/CT was cost effective [36, 37, 39]. Otherwise, the use of PET/CT was not cost effective [34], or only cost effective in a subgroup of patients [33].

In contrast, a cost-effectiveness analysis assessing the cost per true-positive test result was performed in two studies [35, 38]. One study performed a cost-effectiveness analysis in which the cost per therapy procedure was assessed [32]. All three studies were based on decision-tree structures. Simulation time horizons longer than 1 year and discount rates for future costs and health outcomes were not reported. Moreover, no WTP thresholds were reported. Consequently, no clear conclusions regarding cost effectiveness were reported. Therefore, the cost-effectiveness outcomes of these studies could not answer the question if PET/CT would be cost effective.

3.6 Probabilistic and VOI Analysis

Two studies only reported a deterministic analysis [32, 35]. The other studies reported a probabilistic analysis [33, 34, 36–39], where it is essential to report which parameters were included in the analysis and how uncertainty surrounding the point estimates was quantified [41]. A one-way sensitivity analysis was reported in all studies except for one study [32]. All parameters were typically included in the sensitivity analysis. Five studies performed additional scenario analyses adjusting disease-related model parameters (e.g. survival-related probabilities) [33, 34, 36, 37, 39]. Six studies performed additional scenario analyses adjusting pathway-related model parameters (e.g. diagnostic accuracy estimators) [34–39]. Noteworthy, one study reported a threshold analysis in which quality-of-life estimates for disease progression and specific therapy cycles were systematically decreased [33]. No study explored alternative model structures in additional scenario analyses.

Five studies included cost-effectiveness planes depicting the distribution of simulated point estimates [33, 34, 36, 38, 39]. Five studies also included cost-effectiveness acceptability curves [33, 34, 36, 37, 39]. Moreover, two studies performed a VOI analysis by calculating the expected value of perfect information (EVPI) [34, 37]. One study described that additional research regarding the use of PET/CT in the follow-up of non-small cell lung cancer has great value and outweighs the associated costs: an EVPI of €282 per person and a Dutch population EVPI of €423 million was estimated given a WTP threshold of €80,000 per quality-adjusted life-year [37]. Another study declared that additional research regarding the use of PET/CT in the follow-up of cervical cancer has no value: an UK population EVPI of zero was estimated as the probability of PET/CT to be cost effective was zero for all considered relevant WTP thresholds [34].

3.7 CHEERS and ECOBIAS Assessments

The CHEERS checklist was used to assess the methodological quality of the included studies. The extent to which individual items from the CHEERS checklist were reported appropriately is provided in the ESM. Items that were commonly reported inadequately (or missing) were (i) the valuation of outcomes, (ii) the rationale and description of the model, and (iii) analytics and assumptions made. Hence, a description of the methods used to measure and value outcomes was often missing [32, 33, 35–38]. Additionally, the detail of the models and why these model structures were chosen were not described [32, 33, 35, 37, 38]. Finally, methods for analysing or statistically transforming data as well as any extrapolation methods were commonly not (adequately) described [32, 34–39].

The ECOBIAS checklist was used to assess the risk of bias of the included studies. The extent to which individual items from the ECOBIAS checklist were reported appropriately is provided in the ESM. In line with the CHEERS assessment, an invalid valuation bias was often observed

Table 3 Study-level overview of the reported model characteristics, evidence used and health economic analysis performed

Study	Evidence source	Level of evidence	Modelling technique (level)	Cycle length	Type of analysis (health outcome)	QoL questionnaire	Discount rates	WTP threshold	ICER (of the most cost-effective strategy)	Probabilistic analysis?	Outcomes and visualisation	PET/CT considered cost effective?	Software used
Van Loon (2010) [37]	Prospective study	IPD	STM (cohort)	6 months	CUA (QALY)	Not reported	[1] Health 1.5%, [2] costs 4%	€80,000/QALY	For PET/CT, the ICER was €69,086 per QALY compared to conventional imaging	Yes	[1] CEAC, [2] EVPI	Yes	Not reported
Rabalais (2012) [32]	Literature search	Aggregated data	DT	N/A	CEA (therapy efficacy)	Not reported	Not reported	Not reported	For PET/CT, the costs were $14,492 to achieve optimal therapy efficacy	No	N/A	Yes	Not reported
Auguste & Meads (2014) [34]	SLR	Aggregated data	STM (cohort)	3 months	CUA (QALY)	EuroQol 5D	[1] Health 3.5%, [2] costs 3.5%	£20,000–£30,000/QALY	For PET/CT, the ICER was over £1 million per QALY compared to conventional imaging	Yes	[1] CE plane, [2] CEAC, [3] EVPI	No	TreeAge Pro 2011
Greuter (2017) [35]	Prospective study	IPD	DT	N/A	CEA (TP)	Not reported	Not reported	Not reported	For PET/CT, the ICER was €927 per TP compared to using no imaging	No	N/A	Yes	Not reported

Table 3 (continued)

Study	Evidence source	Level of evidence	Modelling technique (level)	Cycle length	Type of analysis (health outcome)	QoL questionnaire	Discount rates	WTP threshold	ICER (of the most cost-effective strategy)	Probabilistic analysis?	Outcomes and visualisation	PET/CT considered cost effective?	Software used
Smith & Mehanna (2017) [39]	RCT	IPD	STM (cohort)	1 month	CUA (QALY)	EuroQol 5D	[1] Health 3.5%, [2] costs 3.5%	£20,000/QALY	For PET/CT, an average per-person lifetime cost saving of £1485 and an additional 0.13 QALYs was calculated	Yes	[1] CE plane, [2] CEAC	Yes	SAS 9.3
Fu (2021) [36]	RCT	IPD	STM (cohort)	3 months	CUA (QALY)	Not reported	[1] Health 1.5%, [2] costs 1.5%	CAD $50,000/QALY	For PET/CT, the ICER was CAD $8193 per QALY compared to conventional imaging	Yes	[1] CE plane, [2] CEAC	Yes	TreeAge Pro 2018

Table 3 (continued)

Study	Evidence source	Level of evidence	Modelling technique (level)	Cycle length	Type of analysis (health outcome)	QoL questionnaire	Discount rates	WTP threshold	ICER (of the most cost-effective strategy)	Probabilistic analysis?	Outcomes and visualisation	PET/CT considered cost effective?	Software used
Greuter (2022) [33]	Population-based registry	IPD	PSM	3 weeks	CUA (QALY)	EORTC QLQ-C30	[1] Health 1.5%, [2] costs 4%	€50,000/QALY	Therapy dropout is not cost effective for positive interim PET/CT (iNMB = −€5281) but is cost effective for negative interim PET/CT (iNMB = €5449)	Yes	[1] CE plane, [2] CEAC	Depending on subgroup	Not reported
Rosen (2022) [38]	Literature search	Aggregated data	DT	N/A	CEA (TP)	Not reported	Not reported	Not reported	For interim PET, the ICER was €4397 per TP compared to conventional imaging	Yes	[1] CE plane	Yes	R (version not reported)

CAD Canadian, *CEA* cost-effectiveness analysis, *CEAC* cost-effectiveness acceptability curve, *CE plane* cost-effectiveness plane, *CT* computed tomography, *CUA* cost-utility analysis, *DT* decision tree, *EVPI* expected value of perfect information, *ICER* incremental cost-effectiveness ratio, *iNMB* incremental net monetary benefit, *IPD* individual patient data, *N/A* not applicable, *PET* positron emission tomography, *PSM* partitioned survival model, *QALY* quality-adjusted life-year, *QoL* quality of life, *RCT* randomised controlled trial, *SLR* systematic literature review, *STM* state-transition model, *WTP* willingness-to-pay

[32, 33, 35–38]. Additionally, a limited sensitivity bias was commonly observed (i.e. the four principles of uncertainty—methodological, structural, heterogeneity, parameter — were not reported in sufficient detail) [32, 35–38]. Furthermore, the model-specific aspects of bias items that were most often inadequately reported (or missing) were (i) structural assumption bias [32, 33, 35–38], (ii) wrong model bias [32, 33, 35, 37, 38], (iii) bias related to baseline data [32, 35–38] and (iv) bias related to diagnostic effects [32–34, 37, 38].

4 Discussion

This systematic literature review evaluated cost-effectiveness studies in which PET/CT was compared to morphological imaging (CT or MRI) for therapy response monitoring. This study aimed to identify methodological strengths and weaknesses, including a systematic assessment of the model structure and assumptions made by the authors. A summary of the key recommendations made throughout the discussion is provided in Table 4.

4.1 Imaging Statistics

First of all, all studies addressed radiological response to chemotherapy. Response to immunotherapy was not studied. As treatment with immunotherapy becomes an increasing part of daily oncology practice, it is expected that cost-effectiveness studies will increasingly focus on response to immunotherapy. Particularly for immunotherapy, in which therapy costs are estimated at €100,000 per line of therapy per patient [42], (early) imaging and corresponding therapy dropouts might be easily cost saving. However, correctly estimating the survival consequences of (early) imaging is expected to be challenging because of the hazards of extrapolating from the limited survival evidence generally available [43–48].

Assuming that imaging tests have perfect sensitivity and specificity is unrealistic. Therefore, the health and economic impact of false imaging classifications should be reflected in cost-effectiveness models. A histology-proven tumour response is considered the highest level of evidence but is generally absent and may also suffer from selective sampling errors. The lack of a gold standard test to which imaging classifications could be compared poses a huge challenge and the subsequent estimation of survival benefits thanks to improved imaging remains highly uncertain. Because of the lack of evidence, expert opinion was regularly used to parametrise the simulation models with imaging characteristics, effect and survival parameters. It is also worth noting that survival outcomes highly depend on contextual factors including the availability of therapies and their effectiveness [14, 15]. We suggest that future simulation models include a comprehensive set of scenarios exploring the impact of a plausible range of potential (but marginal) survival benefits.

4.2 Cohort-Based Models

Decision-tree structures were used in multiple studies to reflect imaging results and the subsequent selection or avoidance of therapy [32, 35, 38]. Decision-tree structures are simplistic and often not suitable or realistic to simulate long-term cost and health outcomes [17, 18]. Alternatively, one model was defined following a traditional three-health-state model structure with progression-free, progression and death states [33]. If the intention is

Table 4 Summary of key recommendations

Section	Summary of key recommendations
Imaging statistics	The health and economic impact of false imaging classifications should be reflected in cost-effectiveness models. Models should include a comprehensive set of scenarios exploring the impact of a plausible range of health consequences
Cohort-based models	Both state-transition models and partitioned survival models are preferred over decision-tree structures. Moreover, state-transition models with sufficient health states are preferred over partitioned survival models. Cost-effectiveness models are encouraged to apply a lifetime simulation horizon when feasible
Patient-level models	Cost-effectiveness models that account for time to events would be valuable. Using more advanced modelling methods may not necessarily result in more complex models, but ensures face validity of the models (which can support clinical decision making in oncology)
Individual patient data	Independent of the selected modelling technique (i.e. cohort modelling or patient-level modelling), individual patient data should be used whenever available
Value of information analysis	In line with the Dutch guideline for performing health economic evaluations in 2024, future cost-effectiveness studies should include a value of information analysis, exploring at least the expected value of perfect information and the expected value of partial perfect information
Validation techniques	Cost-effectiveness models should be tested for external and predictive validation when feasible, in addition to the verification of model parameters, face validity and cross validity testing. Furthermore, models should make better use of existing frameworks to report model validation efforts appropriately

that future models should reflect tumour response over multiple lines of therapy rather than per individual therapy line, adopting a three health-state structure becomes problematic [17, 18]. In a three health-state model, the progression-free state typically represents the therapy line for which response is monitored and for which cost-effectiveness estimates are accumulated over discrete time cycles. Although therapy is not necessarily provided until disease progression, this assumption is reasonable and generally made.

Considering later lines of therapies and long-term cost and health effects requires an updated model structure with robust assumptions and an increased number of health states [49]. For example, in the 12th health-state model [34], tumour response-dependent long-term therapy effects and survival outcomes were considered. Health states distinguished between symptomatic and asymptomatic disease progression. In the nine-health-state model [36], tumour response-dependent long-term therapy effects and survival outcomes were also considered. Health states distinguished between surgery, salvage therapy and best supportive care. These studies showed that, for the assumption of homogenous patients within health states [50], these cohort models require a large number of health states to reflect patient heterogeneity appropriately. Alternatively, to keep the number of health states acceptable for simulation, cohort models can be detailed regarding subgroup characteristics and be specific to the impact of a decision on those subgroups [49].

Complexity is further increased when response to therapy is modelled at discrete points in time and adverse events are modelled continuously in time, both capturing their impact on health outcomes. State-transition models require additional health states and robust assumptions [49]. Therefore, adverse events were rarely included as separate events but considered as part of utility and cost estimation [34, 36, 37, 39]. Alternatively, PSM might be performed. There is, however, a need for more guidance on PSM and how it compares with other modelling approaches [51, 52]. For future simulation models, both STM and PSM are preferred over decision trees. Moreover, STM with sufficient health states are preferred over PSM.

Independent of the modelling technique, it is imperative that the simulation time horizon is long enough to include all cost and effect consequences. In oncology, this implies a lifetime simulation horizon, as selecting a shorter time horizon might result in a biased view of health effects and costs [53, 54]. Therefore, we encourage future modelling studies to apply a lifetime simulation horizon.

4.3 Patient-Level Models

A variety of published modelling approaches was identified to assess the cost effectiveness of therapy response monitoring with PET/CT. All publications were limited to a discrete time approach, with the most common being STM [34, 36, 37, 39]. Patient-level modelling approaches that account for time to events were absent. This was surprising given the observed heterogeneity in therapy response events and the corresponding risk of bias, especially, when few events are observed per time cycle [55]. The continuing personalisation of therapy pathways highlights the need for use of more advanced and patient-level modelling methods to accurately represent the complex context of clinical practice and, therefore, to be meaningful to support decision making [22]. It should be denoted that independent of the modelling structure, a sufficient number of individuals must be modelled and guidance for model development is less extensive [49, 56, 57]. Individual modelling requires more computation time but is not limited by the Markovian property [49]. Therefore, using more advanced modelling methods will not necessarily result in more complex models. On the contrary, it will enable models to continue being valuable for what they are intended for, which is supporting decisions in an increasingly complex environment [22].

4.4 IPD

Five studies primarily used IPD gathered from prospective studies [35, 37], RCTs [36, 39], or national registries [33]. Whenever IPD are available, they should be used. It is important to highlight that in health economic evaluations, the availability of IPD can significantly impact the robustness and accuracy of a cost-effectiveness analysis [58]. While the choice between representing patients as a member of a homogenous cohort (Sect. 4.2) or as individuals (Sect. 4.3) is an essential consideration, the presence of IPD is also crucial for ensuring the validity of the modelling results.

In practice, the decision regarding whether to model a series of cohorts or individuals is primarily pragmatic. There are circumstances in which a patient-level model is preferred, for instance, retaining patient characteristics as a continuous value. Representing these changes over time may add complexity, but if the risks of events are determined by such values, a model that represents individuals should be used [17, 18]. Additionally, use individual simulation (agent-based models or discrete-event simulations) when the number of health states required to reflect the decision problem becomes unmanageably large, as individual simulation allows the representation of substantial heterogeneity in characteristics [18, 57]. In general, the more details required for the parameters that predict outcomes, the more reason to select an individual representation [17, 18, 49, 56, 57].

4.5 VOI Analysis

Uncertainty was found surrounding heterogeneity in therapy response rates, imaging statistics and classifications, model structures, and (extrapolated) effect and survival estimations. This uncertainty reflects suboptimal and incomplete evidence. There might be value in reducing uncertainty, through the collection of new evidence, to better inform and improve decision making in therapy response monitoring. A VOI analysis provides a formal assessment of the value of additional research, based on the extent to which the information improves the expected payoffs associated with a decision by reducing uncertainty. This value is compared with the cost of acquiring the information to determine whether it is worthwhile. [59] There are several VOI measures that can be studied, including the EVPI, expected value of partial perfect information, expected value of sample information and expected net benefit of sampling [60]. When PET/CT is not cost effective on the basis of current evidence and VOI measures suggest that research is not worthwhile, then the (potential) usage of PET/CT should be rejected and discontinued. In line with the Dutch guideline for performing health economic evaluations 2024 [61], we suggest future cost-effectiveness studies to explore a VOI analysis including (at least) the EVPI and expected value of partial perfect information.

4.6 Validation Techniques

Model structures and assumptions could only be reviewed based on what is reported. The two main methods for achieving confidence in the value of simulation models are transparency and validation [62]. On the one hand, some validation was evident for all reviewed studies. Reported validation efforts captured limited verification of model parameters, face validity testing wherein experts evaluated the model inputs and outputs, and cross-validity testing wherein results were compared to results from other models. On the other hand, the two strongest forms of validity testing—external validation wherein model results are compared to real-world data and predictive validation wherein model results are compared to prospectively observed events—were not generally reported. When feasible with respect to the decisions being addressed, models should be tested for their events and their prediction of future events [62]. Furthermore, modelling studies should make better use of existing frameworks to report model validation efforts appropriately [28].

4.7 Limitations

This study strictly focussed on the cost effectiveness of therapy response monitoring with PET/CT, rather than (early) radiological therapy response monitoring in general (i.e. including conventional imaging modalities such as CT and MRI), resulting in the low number of inclusions. Furthermore, during the review process, no clear definition for therapy response monitoring was identified. Most studies considered therapy response monitoring when being on treatment as well as the first weeks or even months off treatment [32, 33, 35, 36, 38, 39]. Other studies attached more value to the *intent* of the scan rather than the *timing* of the scan [34, 37]. All studies that considered PET/CT as intended for therapy response monitoring during or within 3 months after stopping therapy were included.

When using CHEERS items to assess the methodological quality and ECOBIAS items to assess the risk of bias of the included studies, two sources of publication bias may have influenced the results. First, relevant details of included studies may have been excluded to meet the manuscript requirements of peer-reviewed scientific journals. In such cases, online supplementary materials and/or full reports within the UK National Institute for Health Research HTA programme were consulted. After consulting supplementary materials and full National Institute for Health Research reports, there was no need to contact study investigators for unreported data or additional details. Second, not all health economic studies performed were published. This might be the case for studies finding the PET/CT strategy to be not cost effective and may limit the generalisability of the extracted information.

Another limitation is that the CHEERS checklist was developed to support the reporting of health economic studies. The checklist was not developed to assess the methodological quality of health economic evaluations. Therefore, we limited ourselves to the extent to which individual items from the checklist were reported. In the absence of a better alternative, the CHEERS checklist has recently been used more frequently for review assessments in oncology [54, 63]. It should be emphasised that an overall scoring mechanism (i.e. assigning weights or summation of different items) was excluded from this study because such a mechanism could mislead readers and reviewers [29].

5 Conclusions

In the current literature, simulation models that assessed the cost effectiveness of therapy response monitoring with PET/CT were based on cohorts of patients rather than individual patients. Therefore, heterogeneity in therapy response events was generally not appropriately reflected. Further research should include more advanced and patient-level modelling approaches to accurately represent the complex context of daily clinical practice and, therefore, to be meaningful to support decision making.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40273-024-01447-y.

Declarations

Funding No funds, grants or other support were received during the preparation of this article.

Conflicts of interest/competing interests Sietse van Mossel, Rafael Emilio de Feria Cardet, Lioe-Fee de Geus-Oei, Dennis Vriens, Hendrik Koffijberg and Sopany Saing have no relevant financial or non-financial interests to disclose.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Availability of data and material The datasets generated and/or analysed during the current study are available in the supplementary information.

Code availability Not applicable.

Authors' contributions All authors contributed to the study's conception and design. Material preparation, data collection and analysis were performed by SvM and RdFC. The first draft of the manuscript was written by SvM and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50:122S-S150.
2. Langer A. A systematic review of PET and PET/CT in oncology: a way to personalize cancer treatment in a cost-effective manner? BMC Health Serv Res. 2010;10:823.
3. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New Response Evaluation Criteria in Solid Tumours: revised RECIST guideline (version 1. 1). Eur J Cancer. 2009;45:228–47.
4. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbé C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009;15:7412–20.
5. Lopci E, Hicks RJ, Dimitrakopoulou-Strauss A, Dercle L, Iravani A, Seban RD, et al. Joint EANM/SNMMI/ANZSNM practice guidelines/procedure standards on recommended use of [18F] FDG PET/CT imaging during immunomodulatory treatments in patients with solid tumors version 1.0. Eur J Nucl Med Mol Imaging. 2022;49:2323–41.
6. Salaün PY, Abgral R, Malard O, Querellou-Lefranc S, Quere G, Wartski M, et al. Good clinical practice recommendations for the use of PET/CT in oncology. Eur J Nucl Med Mol Imaging. 2020;47:28–50.
7. Fanti S, Goffin K, Hadaschik BA, Herrmann K, Maurer T, MacLennan S, et al. Consensus statements on PSMA PET/CT response assessment criteria in prostate cancer. Eur J Nucl Med Mol Imaging. 2021;48:469–76.
8. Jadvar H, Colletti PM, Delgado-Bolton R, Esposito G, Krause BJ, Iagaru AH, et al. Appropriate use criteria for 18F-FDG PET/CT in restaging and treatment response assessment of malignant disease. J Nucl Med. 2017;58:2026–37.
9. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. Review iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 2017;18:143–52.
10. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328–54.
11. Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. BMJ. 2006;332:1089–92.
12. Buck AK, Herrmann K, Stargardt T, Dechow T, Krause BJ, Schreyögg J. Economic evaluation of PET and PET/CT in oncology: evidence and methodologic approaches. J Nucl Med. 2010;51:401–12.
13. Facey F, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. Health Technol Assess. 2007;11:iii–iv, xi–267.
14. Kip MMA, IJzerman MJ, Henriksson M, Merlin T, Weinstein MC, Phelps CE, et al. Toward alignment in the reporting of economic evaluations of diagnostic tests and biomarkers: the AGREEDT Checklist. Med Decis Mak. 2018;38:778–88.
15. van der Pol S, Rojas Garcia P, Antoñanzas Villar F, Postma MJ, van Asselt ADI. Health-economic analyses of diagnostics: guidance on design and reporting. Pharmacoeconomics. 2021;39:1355–63.
16. Koffijberg H, Van Zaane B, Moons KG. From accuracy to patient outcome and cost-effectiveness evaluations of diagnostic tests and biomarkers: an exemplary modelling study. BMC Med Res Methodol. 2013;13:12.
17. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices: overview. A report of the ISPOR-SMDM modeling good research practices task force-1. Med Decis Mak. 2012;32:667–77.
18. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force-2. Med Decis Mak. 2012;32:678–89.
19. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24:355–71.
20. Briggs A, Claxton K, Sculpher M. Decision modelling for health economic evaluation. 1st ed. Oxford: Oxford University Press; 2006.
21. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. Oxford: Oxford University Press; 2015.

22. Degeling K, Koffijberg H, IJzerman MJ. A systematic review and checklist presenting the main challenges for health economic modeling in personalized medicine: towards implementing patient-level models. Expert Rev Pharmacoecon Outcomes Res. 2017;17:17–25.
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372: n71.
24. Beyer T, Townsend DW, Brun T, Kinahan PE, Charron M, Roddy R, et al. A combined PET/CT scanner for clinical oncology. J Nucl Med. 2000;41:1369–79.
25. Kutylowski J. DeepL Translate. Cologne: Deepl SE; 2017. https://www.deepl.com/translator. Accessed 22 Oct 2024.
26. Veritas Health Innovation. Covidence systematic review software. Melbourne (VIC); 2023. www.covidence.org. Accessed 22 Oct 2024.
27. Pieper D, Puljak L. Language restrictions in systematic reviews should not be imposed in the search strategy but in the eligibility criteria if necessary. J Clin Epidemiol. 2021;132:146–7.
28. Vemer P, Corro Ramos I, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics. 2016;34:349–61.
29. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. Pharmacoeconomics. 2022;40:601–9.
30. Adarkwah CC, van Gils PF, Hiligsmann M, Evers SMAA. Risk of bias in model-based economic evaluations: the ECOBIAS checklist. Expert Rev Pharmacoecon Outcomes Res. 2016;16:513–23.
31. Mehanna H, McConkey CC, Rahman JK, Wong WL, Smith AF, Nutting C, et al. PET-NECK: a multicentre randomised phase III non-inferiority trial comparing a positron emission tomography-computerised tomography-guided watch-and-wait policy with planned neck dissection in the management of locally advanced (N2/N3) nodal metastases in patients with squamous cell head and neck cancer. Health Technol Assess. 2017;21:1–122.
32. Rabalais A, Walvekar RR, Johnson JT, Smith KJ. A cost-effectiveness analysis of positron emission tomography-computed tomography surveillance versus up-front neck dissection for management of the neck for N2 disease after chemoradiotherapy. Laryngoscope. 2012;122:311–4.
33. Greuter MJE, Eertink JJ, Jongeneel G, Dührsen U, Hüttmann A, Schmitz C, et al. Cost-effectiveness of shortening treatment duration based on interim PET outcome in patients with diffuse large B-cell lymphoma. Clin Lymphoma Myeloma Leuk. 2022;22:382–92.
34. Auguste P, Barton P, Meads C, Davenport C, Małysiak S, Kowalska M, et al. Evaluating PET-CT in routine surveillance and follow-up after treatment for cervical cancer: a cost-effectiveness analysis. BJOG. 2014;121:464–76.
35. Greuter MJE, Schouten CS, Castelijns JA, de Graaf P, Comans EFI, Hoekstra OS, et al. Cost-effectiveness of response evaluation after chemoradiation in patients with advanced oropharyngeal cancer using 18F-FDG-PET-CT and/or diffusion-weighted MRI. BMC Cancer. 2017;17:256.
36. Fu TS, Scheffler P, Forner D, Noel CW, Huang SH, Gilbert RW, et al. A cost-utility analysis comparing CT surveillance, PET-CT surveillance, and planned postradiation neck dissection for advanced nodal HPV-positive oropharyngeal cancer. Cancer. 2021;127:3372–80.
37. van Loon J, Grutters JPC, Wanders R, Boersma L, Dingemans AMC, Bootsma G, et al. 18FDG-PET-CT in the follow-up of non-small cell lung cancer patients after radical radiotherapy with or without chemotherapy: an economic evaluation. Eur J Cancer. 2010;46:110–9.
38. Rosen J, Ceccon G, Bauer EK, Werner JM, Tscherpel C, Dunkl V, et al. Cost effectiveness of 18F-FET PET for early treatment response assessment in glioma patients after adjuvant temozolomide chemotherapy. J Nucl Med. 2022;63:1677–82.
39. Smith AF, Hall PS, Hulme CT, Dunn JA, McConkey CC, Rahman JK, et al. Cost-effectiveness analysis of PET-CT-guided management for locally advanced head and neck cancer. Eur J Cancer. 2017;85:6–14.
40. Meads C, Auguste P, Davenport C, Małysiak S, Sundar S, Kowalska M, et al. Positron emission tomography/computerised tomography imaging in detecting and managing recurrent cervical cancer: systematic review of evidence, elicitation of subjective probabilities and economic modeling. Health Technol Assess. 2013;17:7–144.
41. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM modeling good research practices task force-6. Value Health. 2012;15:835–42.
42. Franken MG, Leeneman B, Aarts MJB, van Akkooi ACJ, van den Berkmortel FWPJ, Boers-Sonderen MJ, et al. Trends in survival and costs in metastatic melanoma in the era of novel targeted and immunotherapeutic drugs. ESMO Open. 2021;6: 100320.
43. Kearns B, Stevens J, Ren S, Brennan A. How uncertain is the survival extrapolation? A study of the impact of different parametric survival models on extrapolated uncertainty about hazard functions, lifetime mean survival and cost effectiveness. Pharmacoeconomics. 2020;38:193–204.
44. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, et al. Extrapolating survival from randomized trials using external data: a review of methods. Med Decis Mak. 2017;37:377–90.
45. Davies C, Briggs A, Lorgelly P, Garellick G, Malchau H. The, "hazards" of extrapolating survival curves. Med Decis Mak. 2013;33:369–80.
46. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. Med Decis Mak. 2014;34:343–51.
47. Vickers A. An evaluation of survival curve extrapolation techniques using long-term observational cancer data. Med Decis Mak. 2019;39:926–38.
48. Ouwens MJNM, Mukhopadhyay P, Zhang Y, Huang M, Latimer N, Briggs A. Estimating lifetime benefits associated with immuno-oncology therapies: challenges and approaches for overall survival extrapolations. Pharmacoeconomics. 2019;37:1129–38.
49. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3. Value Health. 2012;15:812–20.
50. Gibson EJ, Begum N, Koblbauer I, Dranitsaris G, Liew D, McEwan P, et al. Cohort versus patient level simulation for the economic evaluation of single versus combination immuno-oncology therapies in metastatic melanoma. J Med Econ. 2019;22:531–44.
51. Woods BS, Sideris E, Palmer S, Latimer N, Soares M. Partitioned survival and state transition models for healthcare decision making in oncology: where are we now? Value Health. 2020;23:1613–21.
52. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of survival probabilities for use in cost-effectiveness analyses: a comparison of a multi-state modeling survival analysis approach with partitioned survival and Markov decision-analytic modeling. Med Decis Mak. 2017;37:427–39.
53. Bullement A, Cranmer HL, Shields GE. A review of recent decision-analytic models used to evaluate the economic value of cancer treatments. Appl Health Econ Health Policy. 2019;17:771–80.

54. Degeling K, Vu M, Koffijberg H, Wong HL, Koopman M, Gibbs P, et al. Health economic models for metastatic colorectal cancer: a methodological review. Pharmacoeconomics. 2020;38:683–713.
55. Degeling K, Franken MD, May AM, van Oijen MGH, Koopman M, Punt CJA, et al. Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients. Cancer Epidemiol. 2018;57:60–7.
56. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM modeling good research practices task force-4. Value Health. 2012;15:821–7.
57. Caro JJ, Möller J. Advantages and disadvantages of discrete-event simulation for health economic analyses. Expert Rev Pharmacoecon Outcomes Res. 2016;16:327–9.
58. Bowrin K, Briere JB, Levy P, Millier A, Clay E, Toumi M. Cost-effectiveness analyses using real-world data: an overview of the literature. J Med Econ. 2019;22:545–53.
59. Fenwick E, Steuten L, Knies S, Ghabri S, Basu A, Murray JF, et al. Value of information analysis for research decisions: an introduction: report 1 of the ISPOR value of information analysis emerging good practices task force. Value Health. 2020;23:139–50.
60. Rothery C, Strong M, Koffijberg H (Erik), Basu A, Ghabri S, Knies S, et al. Value of information analytical methods: report 2 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value Health. 2020;23:277–86.
61. Dutch Healthcare Institute. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidzorg. https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg. Accessed 5 Feb 2024.
62. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. Value Health. 2012;15:843–50.
63. Behr CM, Oude Wolcherink MJ, Ijzerman MJ, Vliegenthart R, Koffijberg H, et al. Population-based screening using low-dose chest computed tomography: a systematic review of health economic evaluations. Pharmacoeconomics. 2023;41:1–17.

Authors and Affiliations

Sietse van Mossel[1,2] · Rafael Emilio de Feria Cardet[3] · Lioe-Fee de Geus-Oei[1,2,4] · Dennis Vriens[5] · Hendrik Koffijberg[6] · Sopany Saing[3,6]

✉ Sietse van Mossel
s.vanmossel-1@utwente.nl

[1] Department of Radiology, Leiden University Medical Centre, Leiden, The Netherlands

[2] Biomedical Photonic Imaging Group, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

[3] Centre for Health Economics Research and Evaluation, University of Technology Sydney, Sydney, NSW, Australia

[4] Radiation Science and Technology, Delft University of Technology, Delft, The Netherlands

[5] Department of Medical Imaging, Radboud University Medical Centre, Nijmegen, The Netherlands

[6] Health Technology and Services Research, University of Twente, Enschede, The Netherlands