Headroom Analysis for Early Economic Evaluation

Applied Health Economics and Health Policy
https://doi.org/10.1007/s40258-022-00774-5

SYSTEMATIC REVIEW

Headroom Analysis for Early Economic Evaluation: A Systematic Review

Esther A. Boudewijns[1] · Thomas M. Otten[2] · Mariam Gobianidze[1] · Bram L. Ramaekers[2] · Onno C. P. van Schayck[1] · Manuela A. Joore[2]

Accepted: 14 November 2022

Abstract

Objectives The headroom analysis is an early economic evaluation that quantifies the highest price at which an intervention may still be cost effective. Currently, there is no comprehensive review on how it is applied. This study investigated the application of the headroom analysis, specifically (1) how the headroom analysis is framed (2) the analytical approach and sources of evidence used, and (3) how expert judgement is used and reported.

Methods A systematic search was conducted in PubMed, Embase, Web of Science, EconLit, and Google Scholar on 28 April 2022. Studies were eligible if they reported an application of the headroom analysis. Data were presented in tabular form and summarised descriptively.

Results We identified 42 relevant papers. The headroom analysis was applied to medicines (29%), diagnostic or screening tests (29%), procedures, programmes and systems (21%), medical devices (19%), and a combined test and device (2%). All studies used model-based analyses, with 40% using simple models and 60% using more comprehensive models. Thirty-three percent of the studies assumed perfect effectiveness of the health technology, while 67% adopted realistic assumptions. Ten percent of the studies calculated an effectiveness-seeking headroom instead of a cost-seeking headroom. Expert judgement was used in 71% of the studies; 23 studies (55%) used expert opinion, 6 studies (14%) used expert elicitation, and 1 study (2%) used both.

Conclusions Because the application of the headroom analysis varies considerably, we recommend its appropriate use and clear reporting of analytical approaches, level of evidence available, and the use of expert judgement.

1 Introduction

Every year, a huge number of new health technologies are being developed, but only a few make it to the market and are cost effective [1]. Amidst finite health-care resources, economic evaluations within a health technology assessment can be helpful to make decisions about ways to spend resources most beneficially [2, 3]. These

Key Points for Decision Makers

Our study revealed that the application and rigor of the headroom analysis ranged widely, from simple models to more comprehensive models, from assumptions of perfect effectiveness to more realistic assumptions, and from calculating maximum costs to minimal effects.

Expert judgements to handle evidence gaps are regularly used, but there is room for improvement in its use and reporting.

The deployment of the headroom analysis and its rigor depend on its aims, the stage of development, and resources available. We provide recommendations to improve the use and reporting of the headroom analysis, including expert judgements.

✉ Esther A. Boudewijns
esther.boudewijns@maastrichtuniversity.nl

[1] Department of Family Medicine, Care and Public Health Research Institute (CAPHRI), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

[2] Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Centre MUMC+/Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands

Published online: 28 December 2022

evaluations often take place after the health technology has been developed to inform reimbursement and usage decisions [4]. In contrast, economic evaluations during the development of health technologies can be useful to provide timely information on investment decisions for further development and research [2, 5, 6]. During the last decades, there has been a growing interest in these early-stage economic evaluations of innovative technologies [7]. The headroom analysis is a type of early economic evaluation, introduced in the context of economic evaluations of health-care technologies by Cosh and colleagues in 2007 [8, 9]. The headroom analysis assesses the maximum price at which a technology is still cost effective [9, 10]. It is calculated by multiplying the effect an innovation might yield (i.e., the room for improvement in current clinical practice or the 'effectiveness gap') by the willingness to pay (WTP) for this effect, plus net savings or minus net costs to the health service [9, 10]. The headroom analysis can be used for pricing policies and to feed into a return-on-investment calculation. As the headroom analysis cannot guarantee future cost effectiveness, it is most useful for ruling out health technologies that are unlikely to be cost effective in future [11]. The rigor of the headroom analysis depends on its aims, the stage of development, and resources available for assessment. As the development and evidence base progress and more resources become available for assessment, the formality and complexity of the headroom analysis may increase [4]. In the early stages, one could assume that a health technology will result in perfect outcomes (e.g., a health technology gives a utility of 1 or is able to prevent all complications). These optimistic assumptions can be supported by the fact that initial development decisions are taken under conditions of greatest uncertainty but entail the least financial commitment [8]. When more information accrues or if experts' judgements about the expected outcomes of using the health technology are identified, a more realistic and extensive model of the additional health benefit and/or costs and savings can be considered. For a user guide of the headroom analysis, we refer to additional file 1 in McAteer and colleagues [12]. The headroom analysis is a valuable and simple tool to aid decision making in an early stage. It might be used by companies to avoid wasteful spending, by researchers to justify public spending, and by funders to make allocation decisions for grants. Currently, there is no comprehensive review on the framing of the analysis, defined as PICOT (population, intervention, comparator, outcomes, time horizon), perspective and development phase of the innovation. Furthermore, there is little information on how the analysis is applied and how expert judgement, which is often used to inform unknown properties of a new technology, is obtained and reported. Therefore, the aim of this study is to investigate the application of the headroom analysis, specifically (1) how the headroom analysis is framed (2) the analytical approach and sources of evidence used, and (3) how expert judgement is used and reported. Based on our findings, we will provide recommendations to improve the use and reporting of the headroom analysis.

2 Methods

This systematic review aimed to assess the application of headroom analysis. The protocol is registered with PROSPERO (CRD42021225601). The study was performed following the methods of systematic reviews proposed by the Cochrane Collaboration and reported using the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) reporting standard [13, 14].

2.1 Search Strategy and Eligibility Criteria

The search strategy was developed together with a medical librarian from Maastricht University. Searches were conducted on 4 January 2021 in PubMed, Embase, Web of Science, EconLit, and Google Scholar. The search was updated on 28 April 2022. Search terms are listed in Supplementary Table 1. Studies were eligible if they reported on an application of the headroom analysis in the context of health care and were peer-reviewed or were (studies in) PhD-theses. A health technology was defined as *an intervention developed to prevent, diagnose or treat medical conditions; promote health; provide rehabilitation; or organise health-care delivery. The intervention can be a test, device, medicine, vaccine, procedure, program or system* [15]. Conference abstracts, presentations, books, discussion papers, draft papers, and patents were excluded. Studies published in Dutch or in English were included and no restrictions were imposed with regard to the publication date. Self-conceived cases that were used to demonstrate the calculation of a headroom analysis were excluded. Studies whose full text was not available after contacting the authors were also excluded.

2.2 Study Selection

Studies were selected in two phases by two independent researchers (EB and MG). First, the title and abstracts were screened to check the eligibility of the studies. This was performed using the Rayyan QCRI web application [16]. The results were compared and non-conformities were discussed. Then, the full texts of the selected studies were reviewed, discrepancies were discussed, and studies were retained if

they met the inclusion criteria. Another researcher (MJ) was consulted in case of discrepancies that could not be resolved through discussions between EB and MG.

2.3 Data Extraction and Synthesis

Data extraction and synthesis was conducted by two independent researchers (EB and TO) and presented in tabular form. We extracted the following characteristics: authors, author affiliations, publication year, journal, country, funding source and role, funding category, and conflict of interest. Furthermore, the framing of the analysis was examined by extracting information on the target population or disease, intervention (i.e., health technology, type of health technology, purported clinical pathway, and purported impact on clinical pathway), comparator, outcomes, time horizon, perspective, WTP threshold, and development phase of the innovation at the time of assessment.

To assess the analytical approach and sources of evidence, we categorised studies into two general groups: first, studies that used a very simple model (i.e., simple calculations based on maximum effects and downstream expenses and savings), and second, studies that used a more comprehensive model-based analysis (decision tree/Markov model or state transition model/discrete event simulation/system dynamic simulation/agent-based modelling and simulation/dynamic transmission modelling/other). We extracted information concerning the sources of evidence informing the model structure, the validation of the model structure, and input parameters, i.e., transition probabilities, costs prices, cost volumes, and effects, to model current care and to model the impact of the new health technology (primary data: assumption/expert judgement/observational studies or real-world data/randomised controlled trails/micro-costing studies/other; and secondary: systematic review or meta-analysis/non-systematic review/literature/population statistics/costing information/other). We also determined whether the headroom analysis assumed perfect effectiveness or aimed to adopt realistic assumptions. Although the headroom analysis was originally developed to assess the maximum additional costs of the new health technology over the comparator, sometimes the formula was re-arranged to calculate the minimal effects needed for the intervention to be cost effective at a given WTP level and given costs. Therefore, we also extracted information on whether the analysis calculated minimal effects or maximum costs. We also extracted main results, including estimated (incremental) effects, costs, and savings and data on additional analyses, including scenario and sensitivity analyses.

Last, we extracted data on ways in which expert judgement was used and reported. We distinguished between expert opinion and expert elicitation according to the definition of Iglesias et al [17]. The first is used to express data qualitatively, while the second is used to express data quantitatively. We scored (yes/partly/no) the use and reporting of expert judgement using the reporting guidelines for the use of expert judgement in model-based economic evaluations [17]. In this reporting guideline expert opinion is focused on a Delphi study, therefore we adjusted the reporting guideline slightly to make it applicable to other forms of expert opinions (Supplementary Table 2a and 2b).

2.4 Risk of Bias

No included studies were subjected to critical appraisal. As the results of individual studies were not analysed in this review, we argue this was not needed to address to study objectives.

3 Results

From a screening of 493 unique studies, 42 studies were included (Fig. 1, Table 1, and Supplementary Figs 1.1–1.4). The included studies were published between 2006 and 2022 (Fig. 2). The use of the headroom analysis seems to have increased over the last years. Studies were conducted in the Netherlands (n = 22; 52%), the UK (n = 15; 36%), Greece (n = 2; 5%), Canada (n = 2; 5%), and Ghana (n = 1; 2%). All studies reported academic affiliations; two studies (5%) also reported affiliations with industry and one study reported affiliations with the government. Funding in 24 studies (57%) was received from governmental or non-for-profit organisations; two studies (5%) reported funding from commercial entities, four studies (10%) reported funding from both governmental or non-profit organisations and commercial entities; 6 studies (14%) reported no funding; and funding was unclear for 6 studies (14%). No conflict of interest was declared for 21 studies (50%); for 13 (31%) this was unclear and 8 (19%) studies reported a conflict of interest.

3.1 Framing of the Headroom Analysis

The framing of headroom analysis varied considerably (Supplementary Table 3 and Supplementary Figs 2.1–2.6). The population or target disease for the new health technology being evaluated was diverse. Twelve interventions (29%) were medicines (including advanced therapy medicinal products and vaccines), 12 (29%) were tests (including biomarkers), 9 (21%) were classified as other (including procedures, programmes, and systems), 8 (19%) were medical devices, and 1 (2%) was both classified as device and test. Most health technologies aimed to impact treatment (n = 19; 45%) and to add-on to the clinical pathway (n = 20; 48%). Usual care was used as a comparator in all

Fig. 1 Study selection

* With a full PhD-thesis considered as article.

** With each relevant PhD-chapter considered as a separate paper.

studies. Thirty-three studies (79%) used Quality Adjusted Life Years (QALYs) as outcome, 6 studies (14%) used clinical outcomes (e.g., hospitalisations reduced), 2 studies (5%) used both QALYs and clinical outcomes – the latter as intermediate outcomes, and 1 study (2%) valued outcomes in monetary units. The time horizon of the studies ranged from one month to a lifetime. A healthcare perspective was most commonly used, while 7 studies (17%) used a societal perspective. The WTP was often dependent on the national threshold and ranged from €4946 to $100,000, with either €20,000 or £30,000 being the most commonly applied WTP thresholds. The development phase of the health technology at the time of assessment was not reported in almost half of the studies (20 studies; 48%). Health technologies were assessed at different points in the lifecycle, including idea screening (5%), concept development (25%) pre-market (8%), and market access (14%) (Supplementary Table 3).

3.2 Analytical Approach and Sources of Evidence

Supplementary Table 4 and Figures 3.1–3.5 display the analytical approach and sources of evidence used.

Seventeen of 42 studies (40%) used simple models with calculations based on the expected maximum effects and downstream expenses and savings. The other studies used a decision tree ($n = 7$; 17%), Markov model/state transition model ($n = 8$; 19%), combination of a decision tree and Markov model/state transition model ($n = 7$; 17%), discrete event simulation ($n = 2$; 5%), or other modelling approaches ($n = 1$; 2%).

Data sources that were used to model current care varied widely, ranging from assumptions and expert judgement to data from randomised controlled trials and systematic reviews. In 18 studies (43%), the impact of the new health technology on model inputs (transition probabilities, costs, savings, and quality of life) was only informed by assumptions. One study (2%) only used expert judgement to inform the impact of the new health technology on model inputs, and seven studies (17%) used both assumptions and expert judgement. In 16 studies (38%), other data sources were included besides assumptions and expert judgement, including observational studies, randomised controlled trials, systematic reviews, non-systematic literature, and costing information. Where these data sources were used, the health technology was already implemented or used [12, 22, 37]; findings from (early) trials, case-control studies, or costing studies were available [21, 26, 28, 40, 44, 53]; costs or effects were based on closely related health technologies

Table 1 Included studies

First author	Year	References
Archibald (Chapter 4)	2015	[18]
Bakker	2020	[19]
Bakker	2021	[20]
Behr	2022	[21]
Brown	2009	[22]
Buisman	2016	[23]
Cao	2013	[24]
Chapman (Chapter 6)	2012	[25]
Chapman (Chapter 8)	2012	[25]
Chrysos (Chapter 6)	2020	[26]
De Graaf	2018	[27]
De Windt	2017	[28]
Francken	2021	[29]
Frempong	2019	[30]
Frempong	2021	[31]
Hummelink	2017	[32]
Huygens	2019	[33]
Huygens	2020	[34]
Kluytmans	2019	[35]
Kluytmans	2019	[36]
Kruyt	2020	[37]
Landry	2022	[38]
Luime	2016	[39]
Major	2019	[40]
Mandavia	2020	[41]
Markiewicz	2016	[42]
McAteer	2007	[10]
McAteer (Chapter 5)	2011	[12]
McAteer (Chapter 6)	2011	[12]
McAteer (Chapter 7)	2011	[12]
Mewes	2017	[43]
Mital	2021	[44]
Postmus	2012	[45]
Thavorn	2020	[46]
Vaidya	2014	[47]
Vallejo-Torres	2011	[48]
Van de Wetering	2012	[49]
Van Nimwegen	2017	[11]
Visser	2022	[50]
Vlek	2020	[51]
Xiong	2006	[52]
Yao	2012	[53]

[19, 20, 22]; or information on parts of the effects or costs associated with the health technology were available [10, 19, 20, 23, 39, 43, 44].

Regarding the effectiveness gap, 14 studies (33%) assumed perfect effectiveness to calculate the headroom and 28 studies (67%) adopted realistic assumptions of additional health benefit and/or downstream savings and costs (Supplementary Table 4).

While the headroom was developed to calculate the maximum cost of the new health technology over the comparator, four studies (10%) only calculated the minimal effects needed for the intervention to be cost effective ('effectiveness-seeking headroom'). These studies already had information about the costs of the intervention [22, 40, 43, 52]. Four studies (10%) calculated both the maximum costs and the minimal effects (Supplementary Table 4) [12, 26, 36, 53].

Scenario and/or sensitivity analyses were conducted in 30 studies (71%), 15 of which (50%) used probabilistic analyses. Value of information analyses were conducted by seven studies. All seven studies considered the uncertainty surrounding the data, upon which the headroom was based.

3.3 Use and Reporting of Expert Judgement

Supplementary Tables 5 and 6 and Supplementary Figure 4.1 display the use and reporting of expert opinion and expert elicitation. Expert judgement was used in 30 of the 42 studies (71%); 23 studies (55%) used expert opinion, 6 studies (14%) used expert elicitation, and 1 study (2%) used both expert opinion and expert elicitation.

Expert opinion: Of the 24 studies (57%) that used expert opinion, 2 studies used different expert panels for different purposes, resulting in 26 use cases of expert opinion being scored. The number of experts was not reported for 14 use cases (54%). For the other 12 use cases, the number ranged from 1 to 26, with a median of 2.5 experts. The expert panel ('sample') was poorly (i.e., scored a 'no' on the reporting guidelines) described for 65% of the use cases in terms of the definition of an expert in the context of the use case and the selection and composition of the panel; it was well described for 1 use case (4%) and for 8 use cases (31%) it was partly described. Expert opinion was used as input for the structure of the model ($n = 12$), for validation of the model structure ($n = 8$), to model current care ($n = 14$), for technology probabilities associated with the new technology ($n = 7$), for cost volumes associated with the new technology ($n = 1$), for cost prices associated with the new technology ($n = 4$), and for effects associated with the new technology ($n = 1$). The research problem for which expert opinion was sought was clearly defined for 35% of the use cases, partly described for 19% of the use cases, and poorly described for 46% of the use cases. The topic and use of expert judgement ('research rationale') was justified for 23% of the use cases, was partly justified for 19% of the use cases, and poorly justified for 58% of the use cases. Six use cases (23%) reported a clear description of the evidence base ('literature review') for the topic of the use case; for four use cases (15%) the evidence base was

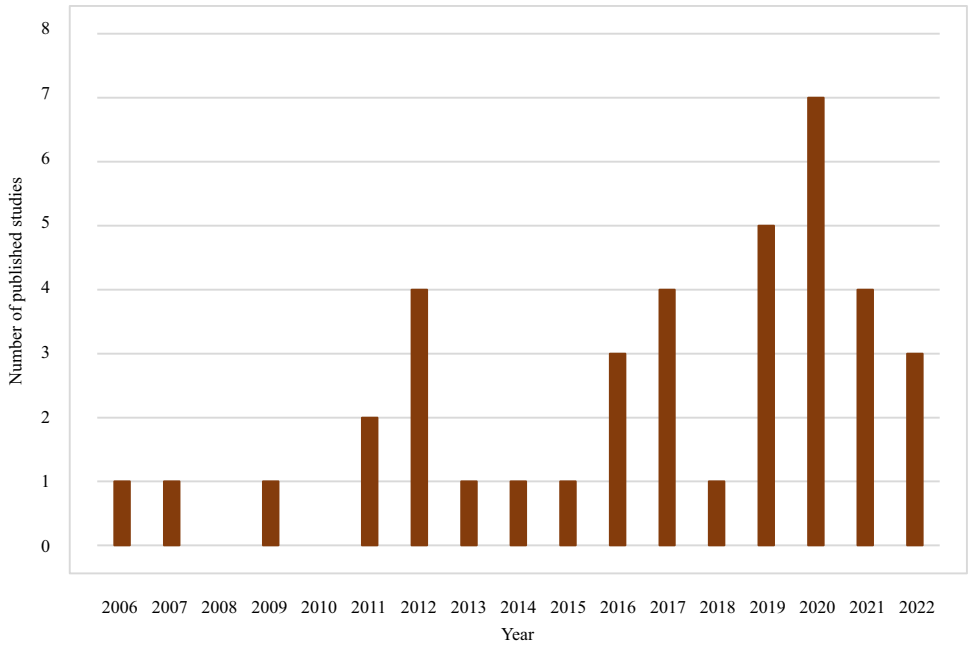

Fig. 2 Number of published studies per year (with a full PhD-thesis considered as article). *The year 2022 includes articles published before April 28

partly described, and 16 use cases (62%) reported poorly on this item. The way in which data were collected and the survey used was poorly reported for 85% of the use cases. Ethical issues and handling of data ('data analysis') were also poorly reported for 96% and 88% of the use cases, respectively. Mixed results were found for the presentation of results, with well-reported results, partly reported results, and poorly reported results for 27%, 35%, and 38% of the use cases, respectively. The interpretation of the results was lacking for 21 use cases (81%).

Expert elicitation: Seven studies (17%) reported the use of expert elicitation. The number of experts was not reported for one study. For the 6 other studies, the number ranged from 2 to 110, with a median of 15 experts. Five studies (71%) did not clearly state the nature of the expert population, with lacking information on the topic of expertise they represented. Expert elicitation was used as input for the structure of the model ($n = 2$), for technology probabilities associated with new technology ($n = 2$), for technology probabilities to model current care ($n = 1$), for cost prices to model current care ($n = 1$), for effects to model current care ($n = 1$) and for effects associated with the new technology ($n = 3$). The need for using an expert elicitation exercise ('research rationale'), the uncertain quantities to be elicited ('research problem'), and measurement type of uncertain quantities were well reported for 57%, 86%, and 86% of the studies, respectively. A clear reference made to a protocol that described the design and conduct of the elicitation exercise ('preparation') was lacking for 5 studies (71%) and was partly described for 2 studies (29%). No study reported on training materials or piloting of the elicitation exercise process. The approach to collect the data ('data collection'), the mode of administering ('administration') and the number and framing of questions ('the exercise') were poorly reported for 43%, 43%, and 57% of the studies, respectively. Well reported information was found for the aggregation method ($n = 4$; 57%), the results ($n = 4$; 57%), and the interpretation of the results ($n = 3$; 43%), while the processes followed to estimate measures of performance for data aggregation and ethical issues were poorly reported by all 7 studies.

4 Discussion

4.1 Summary of Findings

Our systematic review identified 42 studies that applied the headroom analysis in different populations, for different types of health technologies, and at different points in the lifecycle. All studies used model-based analyses, with 40% of the studies using simple models with calculations based on the expected maximum effects and downstream expenses and savings and 60% using more comprehensive models. Thirty-three percent of the studies assumed perfect effectiveness of the health technology, while 67% adopted realistic assumptions. Ten percent of the studies calculated an effectiveness-seeking headroom instead of a cost-seeking headroom. Expert judgement was used in 71% of the studies; 23 studies (55%) used expert opinion, 6 studies (14%) used expert elicitation, and 1 study (2%) used both expert opinion and expert elicitation.

4.2 Previous Research

Several reviews have examined methodologies and tools used in early economic evaluations, including (to a limited extent) the headroom analysis [1–3, 7, 54–56]. They show a wide range of other tools available for early assessment during device development, which should be used appropriately depending on different stages of development or specific goals [1, 54, 55]. Reviews showed that the headroom might be helpful to explore the nature and magnitude of the problem [55] and to think about the added value of the health technology [7]. It was shown that the majority of early economic evaluation methods are considered immature as they lack clear terminology [1, 56]. Several studies mentioned that expert elicitation is of added value in early economic evaluations to inform input parameters where evidence is lacking or is poor [57, 58]. However, it is not routinely applied in early economic evaluations [58] and a universal method for expert elicitation is lacking while different approaches may lead to different results [57]. For example, more guidance is needed on the calibration of experts' judgements to avoid estimates to be neither too optimistic nor too pessimistic [7, 8]. However, it should be noted that conducting an in-depth expert elicitation exercise is not always feasible in the context of an early evaluation, as it might take several months to undertake an elicitation analysis [59].

4.3 Strengths and Limitations

To the best of our knowledge, this is the first systematic review to have provided an in-depth assessment of the application of the headroom analysis. The study included an extensive literature search in four databases and in Google Scholar; the latter allowed us to identify both PhD-theses and academic papers that did not report the analysis in their abstract. There are several limitations to this study. First, we only included scientific, peer-reviewed studies, while headroom analyses will also be conducted by companies that will not publish their analyses because of intellectual property and confidentiality. This will probably have introduced publication bias. Second, for two papers [40, 52], the headroom analysis was not the main topic of the paper, but rather an additional analysis mentioned in the discussion. Consequently, the reporting on the use of the headroom analysis specifically was only minimal for these papers. Third, we only included studies that labelled their analysis as a headroom analysis. However, some studies might have conducted a headroom analysis without using the label. Last, to assess the use and reporting of expert opinion, we had to modify the reporting guidelines for Delphi methods due to a lack of reporting standards for expert opinion [17].

4.4 Recommendations

The headroom analysis is developed to rule-out health technologies that are unlikely to be viable in the future. It can be used by companies to avoid wasteful spending or to convince potential investors of the room for improvement in current practice, for funders to preclude health technologies that do not have the potential to be cost effective, or for researchers to assess whether there is a proof-of-problem and to justify public spending. The deployment of the headroom analysis will depend on its aims, the stage of development, and resources available. Based on the study findings, we recommend the following to improve the use and reporting of the headroom analysis.

We recommend clear labelling of the headroom analysis and improved reporting of funding sources, conflict of interests, and the development phase of the innovation. Our results revealed that the application and rigor of the headroom analysis ranged widely. Therefore, we recommend the use of consistent terminology and clear reporting of the analytical approach and sources of evidence used. This includes whether simple methods or more comprehensive modelling methods were used, whether perfect effectiveness was assumed or realistic assumptions were adopted, and whether minimum effects needed or maximum costs were calculated. Furthermore, because model inputs related to the new health technology are typically uncertain in early stages of development, we recommend giving more attention to reporting of the presence of this uncertainty, and to assessing the impact of these uncertainties using deterministic and probabilistic analyses.

Expert judgement is an important source of data when limited information is available, but our results revealed deficiencies in the use and reporting of expert opinion and expert elicitation. We would like to emphasise that conducting an in-depth expert elicitation exercise is not always feasible due to the costs or time involved. If expert judgement is used, we recommend an increase in the number of experts involved and to better frame the research problem and evidence base to increase the reliability of values identified by expert opinion. For expert elicitation, we recommend making a clear protocol that describes the design and conduct of the exercise, including training materials, and to pilot the exercise process. Regarding reporting of both expert opinion and expert judgement, we recommend that the expert panel is more clearly described, including the topic of expertise they represent and the selection and composition of those involved, as well as the way in which data are collected and ethical issues for the expert sample and research community. We urge for appropriate use of expert judgement (i.e., as well-performed as possible given the available resources) and complete and accurate reporting on its use by means of

reporting guidelines, for example as developed by Iglesias [17].

Furthermore, while our study provided an overview of the application of the headroom analysis, the value and usability of the analysis as perceived by potential users deserves further research [25]. Attention could be paid to unpublished studies conducted by (larger) medical device developers and pharmaceutical companies. Furthermore, the potential role of the headroom analysis to guide decisions regarding the health technology's value proposition and its placement in clinical pathways could be further explored [60].

5 Conclusions

The headroom analysis is a simple analysis to rule out technologies that are unlikely to be cost effective in the future and to understand the existing uncertainty. The use of the headroom analysis to aid decision making in an early stage seems to increase. Our study revealed that the headroom analysis is used in different populations, for different types of health technologies, and at different points in the lifecycle. The application of the analysis ranged widely, from simple models to more comprehensive models, from assumptions of perfect effectiveness to more realistic assumptions, and from calculating maximum costs to minimal effects. The headroom analysis should be conducted appropriately – depending on the aims, stage of development, and resources available – but should always be reported thoroughly. More specifically, clear reporting of analytical approaches and level of evidence available is recommended. Expert judgements to handle evidence gaps are regularly used, but there is room for improvement in its use and reporting. We recommend the use of reporting guidelines.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40258-022-00774-5.

Acknowledgements Gregor Franssen (Maastricht University) provided support in creating the literature search.

Declarations

Funding This study was partly funded by the Netherlands Organisation for Health Research and Development (104006001). The funders has no role in the study design, data collection, analyses, interpretation of the data, or writing of the report.

Conflict of interests The authors declare no conflicts of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Data availability The datasets generated during and/or analysed during the current study are available in the Supplementary Materials. A code is not applicable.

Authors' contributions Electronic searches were performed by EB and MG, together with a medical librarian. Title and abstract screening were conducted by EB and MG. EB and MG conducted the full-text screening. Data extraction and synthesis were conducted by EB and TO, supervised by MJ. EB wrote the first version of the manuscript. TO, MG, BR, OS, and MJ critically reviewed the paper. All authors have read and approved the final version of the manuscript.

References

1. Markiewicz K, Van Til JA, IJzerman MJ. Medical devices early assessment methods: systematic literature review. Int J Technol Assess Health Care. 2014;30(2):137–46.
2. Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: a methodological and empirical review. Int J Technol Assess Health Care. 2008;24(4):465–72.
3. Frempong SN, Sutton AJ, Davenport C, Barton P. Economic evaluation of medical tests at the early phases of development: a systematic review of empirical studies. Expert Rev Pharmacoecon Outcomes Res. 2018;18(1):13–23.
4. Bouttell J, Briggs A, Hawkins N. A toolkit of methods of development-focused health technology assessment. Int J Technol Assess Health Care. 2021;37(1):1–7.
5. Chapman A, Taylor C, Girling A. Early HTA to inform medical device development decisions-the headroom method. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing; 2014: Springer; 2014. p. 1151-4.
6. Love-Koh J. How useful are early economic models?: comment on" Problems and Promises of Health Technologies: The Role of Early Health Economic Modelling". Int J Health Policy Manag. 2020;9(5):215-7.
7. Ijzerman MJ, Koffijberg H, Fenwick E, Krahn M. Emerging use of early health technology assessment in medical product development: a scoping review of the literature. Pharmacoeconomics. 2017;35(7):727–40.
8. Girling A, Lilford R, Cole A, Young T. Headroom approach to device development: Current and future directions. Int J Technol Assess Health Care. 2015;31(5):331–8.
9. Cosh E, Girling A, Lilford R, McAteer H, Young T. Investing in new medical technologies: a decision framework. J Commer Biotechnol. 2007;13(4):263–71.
10. McAteer H, Cosh E, Freeman G, Pandit A, Wood P, Lilford R. Cost-effectiveness analysis at the development phase of a potential health technology: examples based on tissue engineering of bladder and urethra. J Tissue Eng Regen Med. 2007;1(5):343–9.
11. van Nimwegen KJ, Lilford RJ, van der Wilt GJ, Grutters JP. Headroom beyond the quality-adjusted life-year: the case of complex pediatric neurology. Int J Technol Assess Health Care. 2017;33(1):5–10.
12. McAteer H. The use of health economics in the early evaluation of regenerative medicine therapies (Doctoral dissertation). Birmingham, UK: University of Birmingham; 2011.
13. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions. 2nd ed. Chichester: Wiley; 2019.
14. 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. Int J Surg. 2021;88:105906.
15. International Network of Agencies for Health Technology Assessment. HTA Glossary [cited 13-12-2020]. http://htaglossary.net/health+technology.
16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210.
17. Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting guidelines for the use of expert judgement in model-based economic evaluations. Pharmacoeconomics. 2016;34(11):1161–72.
18. Archibald PR. Comparability & reimbursement for the translation of scalable, automated stem cell cultures (Doctoral dissertation). Loughborough, UK: Loughborough University; 2015.
19. Bakker L, Vaporidi K, Aarts J, Redekop W. The potential of real-time analytics to improve care for mechanically ventilated patients in the intensive care unit: an early economic evaluation. Cost Eff Resour Alloc. 2020;18(1):1–11.
20. Bakker L, Aarts J, Uyl-de Groot C, Redekop K. How can we discover the most valuable types of big data and artificial intelligence-based solutions? A methodology for the efficient development of the underlying analytics that improve care. BMC Medical Inform Decis Mak. 2021;21(1):1–12.
21. Behr CM, Koffijberg H, Degeling K, Vliegenthart R, IJzerman MJ. Can we increase efficiency of CT lung cancer screening by combining with CVD and COPD screening? Results of an early economic evaluation. Eur Radiol. 2022;32:3067–75.
22. Brown CA, Lilford RJ. Should the UK government's deep cleaning of hospitals programme have been evaluated? J Infect Prev. 2009;10(4):143–7.
23. Buisman LR, Luime JJ, Oppe M, Hazes JM, Rutten-van Mölken MP. A five-year model to assess the early cost-effectiveness of new diagnostic tests in the early diagnosis of rheumatoid arthritis. Arthritis Res Ther. 2016;18(1):1–12.
24. Cao Q, Postmus D, Hillege HL, Buskens E. Probability elicitation to inform early health economic evaluations of new medical technologies: a case study in heart failure disease management. Value Health. 2013;16(4):529–35.
25. Chapman AM. The use of early economic evaluation to inform medical device decisions: an evaluation of the headroom method (Doctoral dissertation). Birmingham, UK: University of Birmingham; 2012.
26. Chrysos A. Economic analyses of a novel diagnostic device in endocrine disease (Doctoral dissertation). Bristol, UK: University of Bristol; 2020.
27. de Graaf G, Postmus D, Westerink J, Buskens E. The early economic evaluation of novel biomarkers to accelerate their translation into clinical applications. Cost Eff Resour Alloc. 2018;16(1):1–8.
28. de Windt TS, Sorel JC, Vonk LA, Kip MM, Ijzerman MJ, Saris DB. Early health economic modelling of single-stage cartilage repair. Guiding implementation of technologies in regenerative medicine. J Tissue Eng Regen Med. 2017;11(10):2950–9.
29. Francken MF, van Roessel S, Swijnenburg R-J, Erdmann JI, Busch OR, Dijkgraaf MG, et al. Hospital costs of delayed gastric emptying following pancreatoduodenectomy and the financial headroom for novel prophylactic treatment strategies. HPB (Oxford). 2021;23(12):1865–72.
30. Frempong SN, Sutton AJ, Davenport C, Barton P. Early economic evaluation to identify the necessary test characteristics of a new typhoid test to be cost effective in Ghana. PharmacoEconomics-open. 2020;4(1):143–57.
31. Frempong SN, Shinkins B, Howdon D, Messenger M, Neal RD, Sagoo GS. Early economic evaluation of an intervention to improve uptake of the NHS England Diabetes Prevention Programme. Expert Rev Pharmacoecon Outcomes Res. 2022;22(3):417–27.
32. Hummelink S, Gerrits J, Kool LS, Ulrich D, Rovers M, Grutters J. The merits of decision modelling in the earliest stages of the IDEAL framework: an innovative case in DIEP flap breast reconstructions. J Plast Reconstr Aesthet Surg. 2017;70(12):1696–701.
33. Huygens SA, Rutten-van Mölken MP, Noruzi A, Etnel JR, Ramos IC, Bouten CV, et al. What is the potential of tissue-engineered pulmonary valves in children? Ann Thorac Surg. 2019;107(6):1845–53.
34. Huygens SA, Ramos IC, Bouten CV, Kluin J, Chiu ST, Grunkemeier GL, et al. Early cost-utility analysis of tissue-engineered heart valves compared to bioprostheses in the aortic position in elderly patients. Eur J Health Econ. 2020;21(4):557–72.
35. Kluytmans A, Deinum J, Jenniskens K, van Herwaarden AE, Gloerich J, van Gool AJ, et al. Clinical biomarker innovation: when is it worthwhile? Clin Chem Lab. 2019;57(11):1712–20.
36. Kluytmans A, Tummers M, van der Wilt GJ, Grutters J. Early assessment of proof-of-problem to guide health innovation. Value Health. 2019;22(5):601–6.
37. Kruyt IJ, Bours MR, Rovers MM, Hol MK, Rongen J. Economic evaluation of percutaneous titanium implants for bone conduction hearing: a cost-benefit analysis. Otol Neurotol. 2020;41(5):580–8.
38. Landry EC, Scholte M, Su MP, Horstink Y, Mandavia R, Rovers M, et al. Early Health Economic Modelling of Novel Therapeutics in Age-Related Hearing Loss. Front Neurosci. 2022;16:769983.
39. Luime JJ, Buisman LR, Oppe M, Hazes JM, Rutten-van Mölken MP. Cost-effectiveness model for evaluating new diagnostic tests in the evaluation of patients with inflammatory arthritis at risk of having rheumatoid arthritis. Arthritis Care Res. 2016;68(7):927–35.
40. Major RW, Brown C, Shepherd D, Rogers S, Pickering W, Warwick GL, et al. The primary-secondary care partnership to improve outcomes in chronic kidney disease (PSP-CKD) study: a cluster randomized trial in primary care. Am J Nephrol. 2019;30(7):1261–70.
41. Mandavia R, Horstink YM, Grutters JP, Landry E, May C, Rovers M, et al. The potential added value of novel hearing therapeutics: an early health economic model for hearing loss. Otol Neurotol. 2020;41(8):1033–41.
42. Markiewicz K, van Til JA, Steuten LM, IJzerman MJ. Commercial viability of medical devices using Headroom and return on investment calculation. Technol Forecast Soc Change. 2016;112:338-46.
43. Mewes JC, Steuten LM, IJzerman MJ, van Harten WH. A systematic approach for assessing, in the absence of full evidence, whether multicomponent interventions can be more cost-effective than single component interventions. Int J Technol Assess Health Care. 2017;33(4):444–53.
44. Mital S. Early-stage cost-effectiveness analyses of new health care technologies (Doctoral dissertation). Saint John's, Canada: Memorial University of Newfoundland; 2021.
45. Postmus D, de Graaf G, Hillege HL, Steyerberg EW, Buskens E. A method for the early health technology assessment of novel biomarker measurement in primary prevention programs. Stat Med. 2012;31(23):2733–44.
46. Thavorn K, van Katwyk S, Krahn M, Mei SH, Stewart DJ, Fergusson D, et al. Value of mesenchymal stem cell therapy for patients with septic shock: an early health economic evaluation. Int J Technol Assess Health Care. 2020;36(5):525–32.
47. Vaidya A, Joore MA, ten Cate-Hoek AJ, Cate Ht, Severens JL. Cost–effectiveness of risk assessment and tailored treatment for peripheral arterial disease patients. Biomark Med. 2014;8(8):989–99.
48. Vallejo-Torres L, Steuten L, Parkinson B, Girling AJ, Buxton MJ. Integrating health economics into the product development cycle:

a case study of absorbable pins for treating hallux valgus. Med Decis Making. 2011;31(4):596–610.
49. van de Wetering G, Steuten LM, von Birgelen C, Adang EM, IJzerman MJ. Early Bayesian modeling of a potassium lab-on-a-chip for monitoring of heart failure patients at increased risk of hyperkalaemia. Technol Forecast Soc Change. 2012;79(7):1268–79.
50. Visser LA, Folcher M, Delgado Simao C, Gutierrez Arechederra B, Escudero E, Uyl-de Groot CA, et al. The potential cost-effectiveness of a cell-based bioelectronic implantable device delivering interferon-β1a therapy versus injectable interferon-β1a treatment in relapsing-remitting multiple sclerosis. Pharmacoeconomics. 2022;40(1):91–108.
51. Vlek SL, Burm R, Govers TM, Vleugels MP, Tuynman JB, Mijatovic V. Potential value of haptic feedback in minimally invasive surgery for deep endometriosis. Surg Innov. 2020;27(6):623–32.
52. Xiong T, McEvoy K, Morton D, Halligan S, Lilford R. Resources and costs associated with incidental extracolonic findings from CT colonogaphy: a study in a symptomatic population. Br J Radiol. 2006;79(948):948–61.
53. Yao GL, Novielli N, Manaseki-Holland S, Chen Y-F, van der Klink M, Barach P, et al. Evaluation of a predevelopment service delivery intervention: an application to improve clinical handovers. BMJ Qual Saf. 2012;21(Suppl 1):i29–38.
54. Ijzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access. Appl Health Econ Health Policy. 2011;9(5):331–47.
55. Grutters JP, Kluytmans A, van der Wilt GJ, Tummers M. Methods for early assessment of the societal value of health technologies: a scoping review and proposal for classification. Value Health. 2022;25(7):1227–34.
56. Fasterholdt I, Krahn M, Kidholm K, Yderstraede KB, Pedersen KM. Review of early assessment models of innovative medical technologies. Health Policy. 2017;121(8):870–9.
57. Kip MM, Steuten LM, Koffijberg H, IJzerman MJ, Kusters R. Using expert elicitation to estimate the potential impact of improved diagnostic performance of laboratory tests: a case study on rapid discharge of suspected non–ST elevation myocardial infarction patients. J Eval Clin Pract. 2018;24(1):31–41.
58. Steuten LM, Ramsey SD. Improving early cycle economic evaluation of diagnostic technologies. Expert Rev Pharmacoecon Outcomes Res. 2014;14(4):491–8.
59. Bojke L, Soares MFO, Fox A, Jankovic D, Claxton KP, Morton A, et al. Developing a reference protocol for expert elicitation in healthcare decision making. Health Technol Assess. 2021;25(37):1–124.
60. Graziadio S, Winter A, Lendrem BC, Suklan J, Jones WS, Urwin SG, et al. How to ease the pain of taking a diagnostic point of care test to the market: a framework for evidence development. Micromachines. 2020;11(3):291.