RESEARCH

Which value aspects are relevant for the evaluation of medical devices? Exploring stakeholders' views through a Web-Delphi process

Liliana Freitas[1]* , Ana C. L. Vieira[1] , Mónica D. Oliveira[1,2] , Helena Monteiro[3] and Carlos A. Bana e Costa[1,4]

Abstract

Background Implementation and uptake of health technology assessment for evaluating medical devices require including aspects that different stakeholders consider relevant, beyond cost and effectiveness. However, the involvement of stakeholders in sharing their views still needs to be improved.

Objective This article explores the relevance of distinct value aspects for evaluating different types of medical devices according to stakeholders' views.

Methods Thirty-four value aspects collected through literature review and expert validation were the input for a 2-round Web-Delphi process. In the Web-Delphi, a panel of participants from five stakeholders' groups (healthcare professionals, buyers and policymakers, academics, industry, and patients and citizens) judged the relevance of each aspect, by assigning a relevance-level ('Critical', 'Fundamental', 'Complementary', or 'Irrelevant'), for two types of medical devices separately: 'Implantable' and 'In vitro tests based on biomarkers'. Opinions were analysed at the panel and group level, and similarities across devices were identified.

Results One hundred thirty-four participants completed the process. No aspects were considered 'Irrelevant', neither for the panel nor for stakeholder groups, in both types of devices. The panel considered effectiveness and safety-related aspects 'Critical' (e.g., 'Adverse events for the patient'), and costs-related aspects 'Fundamental' (e.g., 'Cost of the medical device'). Several additional aspects not included in existing frameworks' literature, e.g., related to environmental impact and devices' usage by the healthcare professional, were deemed as relevant by the panel. A moderate to substantial agreement across and within groups was observed.

Conclusion Different stakeholders agree on the relevance of including multiple aspects in medical devices' evaluation. This study produces key information to inform the development of frameworks for valuing medical devices, and to guide evidence collection.

Keywords Medical devices, Health technology assessment, Web-Delphi, Health stakeholders, Implantable medical devices, In vitro tests based on biomarkers

*Correspondence:
Liliana Freitas
liliana.freitas@tecnico.ulisboa.pt
Full list of author information is available at the end of the article

BMC

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 2 of 15

Background

Health Technology Assessment (HTA) allows to appraise the relative value of health technologies, to support their introduction and use in healthcare systems, and to inform other healthcare decisions [1–3], such as reimbursement, coverage and pricing [4, 5]. Among health technologies, medicines have been the traditional focus of HTA studies, with HTA agencies often relying on economic evaluation studies (mainly on cost-effectiveness analysis) to inform their decisions [6, 7]. Nevertheless, increasing attention is being given to whether costs and effectiveness are the only relevant aspects to capture health technology's value, namely for evaluating medical devices [1, 3, 4, 8–10].

Medical devices cover a wide range of health technologies, from assistive devices to sophisticated implants, and can facilitate disease prevention, diagnosis, and treatment [11]. The World Health Organization estimates that there are 2 million different medical devices [12]. The diversity and rapid pace of innovation of this industry, together with the recent European regulations [13–15], highlight the need to systematise the HTA process for these types of technologies [4, 9]. When comparing to HTA processes for drugs, researchers highlight undeniable differences that can be brought by the device-specific features [1, 16, 17]. Specifically, the device-operator interaction, the incremental and rapid innovation, or the difficulty in conducting randomised controlled trials to produce high-quality evidence have been features recalled to impact on evidence and on evaluations in practice [3, 7, 18].

Exploring how the evaluation of medical devices is being conducted in Europe, Fuchs et al. [3] interviewed 16 representatives from HTA institutions and observed differences across and within countries in the value aspects considered when evaluating medical devices, with aspects depending on its relevance for evaluators and on the type and purpose of the devices being assessed. These findings relate to a lack of methodological guidance in the area, as pointed out by Ciani et al. [7] when analysing practices across 36 non-European HTA agencies.

Researchers have been attempting to develop evaluation frameworks to standardise and bring guidance and transparency to the evaluation of health technologies, including medical devices [19–21]. Nevertheless, a recent systematic review [22] describing HTA-related value assessment frameworks showed that there is still no consensus on how to define health technology's value with the frameworks differing in the value aspects included. Furthermore, several authors acknowledge that these frameworks fail to consider a wide range and diversity of stakeholders' perspectives [19], with the engagement of patients and public being very limited or missing [22, 23]. The inclusion of stakeholders' views is described as key to enable HTA adoption [24–26], with Mueller et al. [27] concluding – about the importance of involving stakeholders in HTA – that "engaging and consulting stakeholders locally was imperative to understand the context, reduce evidence gaps and address the uncertainties in the evidence, ultimately paving the way for technology adoption" (p.14).

Several of those frameworks have explored the use of multicriteria decision analysis (MCDA) as a frame both to consider multiple evaluation aspects, explicitly and transparently, and to include stakeholders' views within evaluations [2, 24]. Nevertheless, the difficulties in synthesizing key information and in establishing evaluation criteria, and the fact that MCDA modelling usually rely on small numbers of participants, have been pointed out as shortcomings [2, 28]. To overcome these shortcomings, Web-Delphi processes have been used in other health contexts to gather stakeholders' opinions, promote discussions or consensus [29, 30], and to produce information for assisting MCDA modelling [2, 31, 32]. These processes have shown to gather opinions from large and heterogeneous groups, to promote consensus and to generate a collaborative environment, under a low cost format [33]. This article aims to explore the views of different stakeholders regarding the relevance of a set of value aspects for evaluating two distinct types of medical devices, 'Implantable medical devices' and 'In vitro tests based on biomarkers', through a Web-Delphi process. Results from this process can inform HTA processes in practice, discussions about which value aspects are relevant to consider in medical devices evaluation frameworks (including in MCDA-based frameworks), as well as the design of tools for evaluating distinct types of devices.

Methods

Overview of the Web-Delphi process

This study was developed within the scope of the MEDI-VALUE ("Developing HTA tools to consensualise MEDIcal devices' VALUE through multicriteria decision analysis") project [34], a national research project that included as consortium partners the Portuguese national HTA agency (INFARMED) and three leading Portuguese hospitals (Centro Hospitalar Lisboa Norte, Hospital do Espírito Santo de Évora and Instituto Português de Oncologia de Lisboa), and aimed at advancing HTA literature by designing and implementing methods, informed by MCDA, that enable the involvement of a large number of health stakeholders and promote consensus in the structuring and development of sound models for assessing the multidimensional value of medical devices.

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 3 of 15

To inform HTA processes and to inform the building of evaluation models, this study presents a Web-Delphi process designed (see Fig. 1) for gathering different stakeholders' opinions on the relevance of distinct aspects for medical devices' evaluation in terms of their added-value to an alternative comparator. Delphi processes allow to anonymously collect individual opinions in successive rounds presenting, from round 2 onwards, a summary of the opinions given in the previous round to participants [35]. This design allows participants to reflect on their previous opinions and change them, or not, based upon new generated information.

In preparation for this process, the aspects potentially relevant for the evaluation of medical devices were collected from studies using MCDA for medical devices evaluation, identified in the systematic review of Oliveira et al. [2], and by extending the search protocol of that study until October of 2019. From this search, a final list of 34 aspects (described in Additional file 1) was organised, with two MEDI-VALUE experts from INFARMED (the Portuguese HTA regulator) analysing the list completeness, eliminating evident redundancies.

The Web-Delphi was implemented in the Welphi platform [36] and was composed of two rounds. In the first round (12 March-02 May 2020), participants were invited to give their opinion about the relevance of each of the 34 aspects for each of the two types of medical devices, by choosing one level of the four-level qualitative relevance scale presented below. This relevance scale allows not only to identify which aspects are relevant for stakeholders, but also to capture the strength of the relevance, either for the panel or for each stakeholder group. Following the rational of "determinants" [37] and "relevancy" [38] analysis (which is context specific [39]), this information is useful for informing the structuring of a multicriteria value framework, as it helps in screening out non-relevant aspects, and for assisting the building of multicriteria evaluation models, as it differentiates the relevance among aspects. The scale was developed,

tested and validated with five health experts from MEDI-VALUE partners.

- **Critical**: this aspect, beyond fundamental, can, by itself, preclude assessing if the medical device has added value given its alternative.
- **Fundamental**: this aspect must, undoubtedly, be part of the basis of evaluation of the medical device, to assess if it has added value given its alternative.
- **Complementary**: this aspect is not fundamental but, still, it can add something to the value of the medical device given its alternative.
- **Irrelevant**: this aspect must not be part of the basis of evaluation of the medical device; it is inapplicable or irrelative to assess if the medical device has added value given its alternative.

The 34 aspects and the relevance scale were presented to participants in two separate screens: for 'Implantable medical devices' and for 'In vitro tests based on biomarkers' (a therapeutic and a diagnostic type of device, respectively). For each type of medical device, participants could provide general comments and, for each aspect individually, they could select the 'Don't know/ Don't want to answer' option or provide specific comments. Participants could give their opinions for one or both types of medical devices. In the second round (9–31 May 2020), participants had access to similar screens, but then they could access their own previous answers, and the distribution of all first-round answers and comments as feedback, being able to change their answer. After the second round was concluded, participants could access the results and leave comments regarding the results and the process (23 June-14 July 2020). A second study was developed simultaneously, exploring the effect of feeding back the distribution of the answers disaggregated per groups of stakeholders on participants' opinion change. For this, half of the participants were randomly selected to also have access to such information. A manuscript of

Web-Delphi preparation	Web-Delphi process		
Identifying potentially relevant evaluation aspects	Defining the level of relevance of each aspect for medical devices evaluation, according to different groups of stakeholders		
• **Literature review** • **Validation with experts from the Portuguese HTA regulator**	**Round 1** - Stating the level of relevance of each evaluation aspect by each participant - Providing comments	**Round 2** (based on summary round 1) - Confirming/revising the level of relevance of each evaluation aspect by each participant - Providing comments	**Presentation of results** - Providing comments to process and results

Fig. 1 Design of the Web-Delphi process
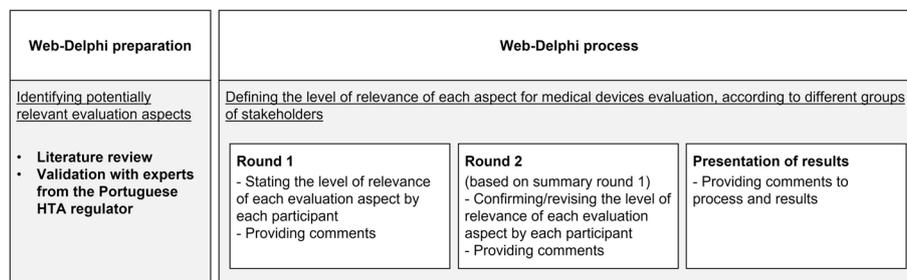
Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 4 of 15

this second study is being prepared, and more information can be provided upon request.

Invited participants
Participants' selection followed a purposive sampling strategy targeting five stakeholders' groups in Portugal: healthcare professionals (doctors, nurses, pharmacists, technicians), buyers and policymakers, academics, industry, and patients and citizens. MEDI-VALUE partners disseminated the study among their networks, explaining the study's aim and the type of participants being recruited. Specifically, selected stakeholders should: (a) have experience in the use, evaluation, selection or acquisition of medical devices, or having an interest in the topic; (b) be available to participate given the timeline planned; (c) not have any conflict of interest preventing their impartial participation. Through this process, a total of 365 stakeholders were identified and invited to participate. As each participant can have several roles in practice and, thus, could belong to more than one group, participants selected, at the beginning of the Web-Delphi, the stakeholder group in the light of which they would give their opinions.

Analysis
Analyses were performed in Microsoft Excel and R software to answer three specific research questions: to explore the relevance of each aspect for the evaluation of the two types of medical devices, for the panel and per stakeholders' groups; to measure the level of agreement within groups; and to explore whether there were aspects with the same relevance for the two types of medical devices.

Relevance of each aspect for the evaluation of the two types of medical devices
For each type of medical device, the answers were analysed at the panel-level and, then, at the stakeholders' groups-level. A panel majority opinion about each aspect's relevance was defined at the end of the process, based upon the absolute majority of answers (more than 50%). Specifically, if there was a relevance level gathering absolute majority, it would be selected to describe the panel majority opinion, otherwise panel majority opinion would be described as 'No majority'. The distribution of answers per group of stakeholders was analysed to draw conclusions on how aligned the groups were with the panel, in each aspect. To complement this analysis, the Kruskal–Wallis H-test [40], with 4 degrees of freedom, was used to explore whether the results were significantly different across the distinct groups. As this test only reveals if there is a difference and does not specify between which groups the differences occur, the post-hoc

Dunn's test was performed, corrected with the Bonferroni test, for the cases in which the Kruskal–Wallis test found significant differences [40]. Differences across stakeholders' groups were considered significant for Dunn's-Bonferroni p-value lower than 0.05.

Agreement within groups of stakeholders for the two types of medical devices
The Gwet's AC2 agreement coefficient [41, 42] was used to determine the inter-rater reliability within each group of stakeholders, with the quadratic weighting scheme [43] used to compute the coefficient. Afterwards, Gwet's coefficients were compared to the Landis and Koch benchmark levels [44] that establish the agreement as poor (for coefficient values < 0), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) or almost perfect (0.81–1), This comparison allowed us to determine the strength of agreement according to the Landis and Koch scale [44]. Despite the high support of this benchmarking scale among researchers [45], Gwet [41] points out that the AC coefficients have a probability distribution and an error margin associated with them, and that the benchmarking approach should account for that uncertainty. Thus, we have also applied Gwet's proposed benchmarking method [41] that determines the agreement level (of a selected benchmarking scale) associated with a 95% confidence. We have adopted the Landis and Koch benchmark scale, due to its finer categorization [41], and considered the 95% cut-off point for the cumulative probabilities. As this benchmarking method considers the standard error of the computed coefficients, the choice of the agreement level can be more conservative than applying Landis and Koch benchmark alone [41].

Common relevance for both types of medical devices
Finally, it was explored whether similar conclusions about the relevance of each aspect could be retrieved for both types of medical devices simultaneously. For that purpose, the panel and the groups majority opinions were compared, and the aspects gathering the same relevance in both types of medical devices were identified.

Results
Web-Delphi's first round was completed by 167 participants, with 134 (80.2%) completing the second round. The distribution of participants per group of stakeholders and the dropout rates are presented in Table 1.

Table 1 also details how many participants concluded the Web-Delphi for 'implantable' devices (127) and for 'in vitro' devices (119). These numbers of participants were

Table 1 Participants in the Web-Delphi process, per group of stakeholders, and dropout rate

Groups of stakeholders	Round 1			Round 2			Dropout (%)
	Implantable medical devices	In vitro tests based on biomarkers	Total	Implantable medical devices	In vitro tests based on biomarkers	Total	
Academics	26	25	27	22	21	23	14,8%
Buyers and policymakers	10	8	10	8	6	8	20,0%
Healthcare professionals (doctors, nurses, pharmacists, technicians)	69	65	74	52	48	57	23,0%
Industry	16	15	17	12	11	13	23,5%
Patients and citizens	39	39	39	33	33	33	15,4%
TOTAL	160	152	167	127	119	134	19,8%

used in the following analyses, as each type of medical of devices is looked at individually.

Relevance of each aspect for the evaluation of each of the two types of medical devices

Table 2 presents the distribution of answers, per type of medical device, of the panel and disaggregated per group of stakeholders. Moreover, it presents the panel majority opinion and the analysis on whether the groups are aligned with the panel, described as the groups majority opinions.

For 'Implantable medical devices' six aspects gathered panel majority on 'Critical', 16 on 'Fundamental', and one on 'Complementary'. Regarding 'In vitro tests based on biomarkers', seven aspects gathered panel majority on 'Critical', 13 on 'Fundamental' and one on 'Complementary'. For both types of medical devices, the remaining aspects had no panel majority opinion. These results suggest that the panel considers all 34 aspects as relevant to evaluate the added value of a medical device for any of these types of devices – as there was no aspects gathering majority on 'Irrelevant' – and, specifically, that the panel agrees in the relevance of a high number of aspects – 23 and 21 aspects for 'implantable' and 'in vitro' devices, respectively.

When analysing the answers per groups of stakeholders, we observe some differences. For 'Implantable medical devices', for four of the six above-mentioned 'Critical' aspects and three of the 11 'Fundamental' aspects, the majority occurred across all stakeholders' groups, with the remaining aspects presenting one to three groups of stakeholders without the same majority opinion. Identically, for 'In vitro tests based on biomarkers', only two of the five 'Critical' aspects and two of the 11 'Fundamental' aspects presented the same majority across all groups. Regarding the aspects with no panel majority opinion, most present a majority on one to three groups of stakeholders, with only two aspects ('Patient-reported

outcomes' and 'Environmental impact of the production and use of the medical device') of 'in vitro' devices not gathering majority in any group.

In all cases, for the aspects with no groups majority opinions, the answers of the groups tend to be around the same two relevance levels (these two levels are highlighted in light grey in Table 2). Exceptions to this are, for example, 'Exposure of the healthcare professional to physical or chemical agents' that presents groups of stakeholders with most answers around 'Critical' and 'Fundamental' and other groups with most answers around 'Fundamental' and 'Complementary' (this happening for both types of medical devices, with a panel majority opinion defined only in 'implantable' devices).

For 'Implantable medical devices' a statistically significant difference (statistical tests results available in Additional file 2) was found between Industry and Patients and citizens in 'Sensitivity and Specificity' ($p = 0.0449$), for which Industry mostly considered 'Critical' and Patients and citizens had a dispersed opinion with no majority, and 'Impact of the disease' ($p = 0.0282$), for which Industry group's opinion gathered no majority and Patients and citizens gathered majority in 'Fundamental'; between Healthcare professionals and Patients and citizens in 'Exposure of the healthcare professional to physical or chemical agents' ($p = 0.0405$), for which Healthcare professionals gathered no majority and Patients and citizens gathered majority in 'Fundamental'; between Academics and Patients and citizens in 'Target population' ($p = 0.0447$) and 'Cost of the medical device (including complementary equipment)' ($p = 0.0266$), in which both groups gathered majority on 'Fundamental' but Academics gathered opinion among 'Critical' and 'Fundamental', and Patients and citizens among 'Fundamental' and 'Complementary'; and between Academics and Healthcare professionals in 'Cost of procedure without the cost of the medical device' ($p = 0.0326$), with both groups gathering majority in 'Fundamental' but Academics

Table 2 Aspects, distribution of answers, panel majority opinion and groups majority opinions

Aspect	Stakeholders' group	IMPLANTABLE MEDICAL DEVICES						IN VITRO TESTS BASED ON BIOMARKERS					
		Critical	Fundamental	Complementary	Irrelevant	Don't know...	Conclusions on the alignment of the distinct groups of stakeholders	Critical	Fundamental	Complementary	Irrelevant	Don't know...	Conclusions on the alignment of the distinct groups of stakeholders
Specific features of the medical device	Overall panel	69%	27%	3%	0%	1%	*Panel majority opinion*: Critical	43%	50%	6%	1%	1%	*Panel majority opinion*: No majority
	Academics	73%	18%	9%	0%	0%	*Groups majority opinions: All Critical*	43%	48%	5%	0%	5%	*Groups majority opinions: Critical for Buyers and policymakers, Fundamental for Patients and citizens. No majority for the remaining groups.*
	Buyers and policymakers	63%	38%	0%	0%	0%		83%	17%	0%	0%	0%	
	Industry	75%	25%	0%	0%	0%		45%	45%	9%	0%	0%	
	Patients and citizens	61%	36%	0%	0%	3%		30%	58%	9%	3%	0%	
	Healthcare professionals	73%	23%	4%	0%	0%		46%	50%	4%	0%	0%	
Technical performance of the medical device	Overall panel	79%	20%	0%	0%	2%	*Panel majority opinion*: Critical	69%	29%	0%	0%	2%	*Panel majority opinion*: Critical
	Academics	86%	14%	0%	0%	0%	*Groups majority opinions: All Critical*	67%	33%	0%	0%	0%	*Groups majority opinions: All Critical*
	Buyers and policymakers	63%	38%	0%	0%	0%		83%	17%	0%	0%	0%	
	Industry	92%	8%	0%	0%	0%		64%	27%	0%	0%	9%	
	Patients and citizens	73%	21%	0%	0%	6%		61%	36%	0%	0%	3%	
	Healthcare professionals	79%	21%	0%	0%	0%		75%	25%	0%	0%	0%	
Regulatory status of the medical device	Overall panel	38%	47%	13%	1%	2%	*Panel majority opinion*: No majority	37%	42%	19%	2%	0%	*Panel majority opinion*: No majority
	Academics	50%	45%	0%	5%	0%	*Groups majority opinions: Critical for Industry, Fundamental for Buyers and policymakers and Patients and citizens. No majority for the remaining groups.*	43%	38%	14%	5%	0%	*Groups majority opinions: Critical for Buyers and policymakers. No majority for the remaining groups.*
	Buyers and policymakers	38%	63%	0%	0%	0%		67%	33%	0%	0%	0%	
	Industry	67%	17%	17%	0%	0%		45%	36%	18%	0%	0%	
	Patients and citizens	24%	55%	15%	0%	6%		24%	45%	27%	3%	0%	
	Healthcare professionals	35%	48%	17%	0%	0%		38%	44%	19%	0%	0%	
Sensitivity and Specificity	Overall panel	45%	46%	5%	2%	2%	*Panel majority opinion*: No majority	75%	24%	2%	0%	0%	*Panel majority opinion*: Critical
	Academics	64%	36%	0%	0%	0%	*Groups majority opinions: Critical for Academics and Industry, Fundamental for Buyers and policymakers. No majority for the remaining groups.*	76%	24%	0%	0%	0%	*Groups majority opinions: All Critical, except Buyers and policymakers (=50%).*
	Buyers and policymakers	13%	75%	0%	13%	0%		50%	50%	0%	0%	0%	
	Industry	75%	25%	0%	0%	0%		100%	0%	0%	0%	0%	
	Patients and citizens	30%	45%	12%	6%	6%		70%	24%	6%	0%	0%	
	Healthcare professionals	44%	50%	4%	0%	2%		75%	25%	0%	0%	0%	
User-friendliness for the healthcare professional	Overall panel	11%	66%	21%	0%	2%	*Panel majority opinion*: Fundamental	14%	55%	27%	3%	1%	*Panel majority opinion*: Fundamental
	Academics	18%	59%	23%	0%	0%	*Groups majority opinions: All Fundamental*	14%	52%	29%	0%	5%	*Groups majority opinions: All Fundamental, except Buyers and policymakers*
	Buyers and policymakers	0%	63%	38%	0%	0%		17%	33%	50%	0%	0%	
	Industry	0%	83%	17%	0%	0%		18%	55%	27%	0%	0%	
	Patients and citizens	6%	64%	24%	0%	6%		3%	55%	36%	6%	0%	
	Healthcare professionals	15%	67%	17%	0%	0%		21%	60%	17%	2%	0%	
Time between procedure and results	Overall panel	12%	49%	35%	2%	2%	*Panel majority opinion*: No majority	10%	71%	18%	1%	0%	*Panel majority opinion*: Fundamental
	Academics	18%	45%	36%	0%	0%	*Groups majority opinions: Fundamental for Patients and citizens. No majority for the remaining groups.*	10%	62%	29%	0%	0%	*Groups majority opinions: All Fundamental*
	Buyers and policymakers	0%	38%	50%	0%	13%		0%	83%	17%	0%	0%	
	Industry	17%	33%	50%	0%	0%		9%	73%	18%	0%	0%	
	Patients and citizens	9%	58%	27%	0%	6%		15%	64%	21%	0%	0%	
	Healthcare professionals	12%	50%	33%	6%	0%		8%	77%	13%	2%	0%	
Need for training of the healthcare professional	Overall panel	11%	54%	31%	2%	2%	*Panel majority opinion*: Fundamental	8%	59%	28%	4%	1%	*Panel majority opinion*: Fundamental
	Academics	14%	55%	32%	0%	0%	*Groups majority opinions: All Fundamental, except Buyers and policymakers (=50%) and Industry*	10%	52%	33%	0%	5%	*Groups majority opinions: All Fundamental, except Buyers and policymakers (=50%) and Industry*
	Buyers and policymakers	0%	50%	50%	0%	0%		0%	50%	50%	0%	0%	
	Industry	33%	25%	42%	0%	0%		18%	45%	27%	9%	0%	
	Patients and citizens	3%	58%	30%	3%	6%		3%	55%	36%	6%	0%	
	Healthcare professionals	12%	60%	25%	4%	0%		10%	69%	17%	4%	0%	
Learning curve of the healthcare professional	Overall panel	5%	44%	45%	4%	2%	*Panel majority opinion*: No majority	6%	39%	52%	3%	1%	*Panel majority opinion*: Complementary
	Academics	9%	45%	45%	0%	0%	*Groups majority opinions: Fundamental for Healthcare Professionals, Complementary for Buyers and policymakers and Patients and citizens. No majority for the remaining groups.*	10%	29%	57%	0%	5%	*Groups majority opinions: All Complementary, except Healthcare professionals*
	Buyers and policymakers	0%	13%	75%	0%	13%		0%	17%	83%	0%	0%	
	Industry	8%	50%	25%	17%	0%		9%	36%	55%	0%	0%	
	Patients and citizens	3%	27%	58%	6%	6%		3%	36%	58%	3%	0%	
	Healthcare professionals	4%	58%	37%	2%	0%		6%	48%	42%	4%	0%	
Exposure of the healthcare professional to physical or chemical agents	Overall panel	26%	50%	18%	4%	2%	*Panel majority opinion*: Fundamental	39%	45%	13%	3%	1%	*Panel majority opinion*: No majority
	Academics	18%	59%	23%	0%	0%	*Groups majority opinions: All Fundamental, except Buyers and policymakers and Healthcare professionals*	24%	52%	19%	0%	5%	*Groups majority opinions: Fundamental for Academics and Patients and citizens. No majority for the remaining groups.*
	Buyers and policymakers	13%	38%	38%	13%	0%		17%	33%	50%	0%	0%	
	Industry	17%	58%	17%	8%	0%		36%	36%	9%	18%	0%	
	Patients and citizens	12%	55%	21%	6%	6%		39%	52%	6%	3%	0%	
	Healthcare professionals	42%	44%	12%	2%	0%		48%	40%	10%	2%	0%	
Workload for the healthcare professional	Overall panel	6%	46%	40%	6%	2%	*Panel majority opinion*: No majority	8%	47%	40%	3%	2%	*Panel majority opinion*: No majority
	Academics	9%	50%	36%	5%	0%	*Groups majority opinions: Fundamental for Healthcare professionals. No majority for the remaining groups.*	10%	62%	24%	0%	5%	*Groups majority opinions: Fundamental for Academics and Healthcare professionals, Complementary for Buyers and policymakers. No majority for the remaining groups.*
	Buyers and policymakers	0%	25%	50%	13%	13%		0%	17%	83%	0%	0%	
	Industry	17%	33%	42%	8%	0%		9%	45%	45%	0%	0%	
	Patients and citizens	3%	45%	39%	6%	6%		9%	33%	45%	9%	3%	
	Healthcare professionals	4%	52%	40%	4%	0%		6%	54%	38%	2%	0%	
Comfort for the patient	Overall panel	26%	62%	10%	0%	2%	*Panel majority opinion*: Fundamental	29%	51%	15%	4%	1%	*Panel majority opinion*: Fundamental
	Academics	36%	45%	18%	0%	0%	*Groups majority opinions: All Fundamental, except Academics.*	33%	52%	10%	0%	5%	*Groups majority opinions: Fundamental for Academics and Healthcare Professionals. No majority for the remaining groups (=50% on Fundamental for Buyers and policymakers).*
	Buyers and policymakers	0%	75%	25%	0%	0%		0%	50%	33%	17%	0%	
	Industry	33%	67%	0%	0%	0%		27%	36%	27%	9%	0%	
	Patients and citizens	30%	55%	9%	0%	6%		30%	48%	12%	9%	0%	
	Healthcare professionals	21%	71%	8%	0%	0%		29%	56%	15%	0%	0%	

Table 2 (continued)

Criterion	Group						Left majority opinion						Right majority opinion
Connectivity	Overall panel	4%	46%	45%	2%	3%	*Panel majority opinion*: No majority	6%	39%	44%	10%	2%	*Panel majority opinion*: No majority
	Academics	14%	32%	36%	9%	9%	*Groups majority opinions:* Fundamental for Industry and Patients and citizens, Complementary for Healthcare Professionals. No majority for the remaining groups.	10%	24%	52%	10%	5%	*Groups majority opinions:* Complementary for Academics. No majority for the remaining groups.
	Buyers and policymakers	0%	50%	50%	0%	0%		0%	50%	17%	33%	0%	
	Industry	8%	67%	25%	0%	0%		9%	36%	45%	9%	0%	
	Patients and citizens	0%	52%	42%	0%	6%		3%	39%	45%	12%	0%	
	Healthcare professionals	2%	42%	54%	2%	0%		6%	44%	42%	6%	2%	
Clinical efficacy and/or effectiveness	Overall panel	80%	18%	0%	0%	2%	*Panel majority opinion*: Critical	76%	20%	1%	1%	2%	*Panel majority opinion*: Critical
	Academics	86%	14%	0%	0%	0%	*Groups majority opinions:* All Critical	76%	14%	0%	0%	10%	*Groups majority opinions:* All Critical
	Buyers and policymakers	75%	25%	0%	0%	0%		100%	0%	0%	0%	0%	
	Industry	100%	0%	0%	0%	0%		91%	9%	0%	0%	0%	
	Patients and citizens	73%	21%	0%	0%	6%		70%	24%	3%	3%	0%	
	Healthcare professionals	79%	21%	0%	0%	0%		75%	25%	0%	0%	0%	
Risk analysis	Overall panel	65%	29%	4%	0%	2%	*Panel majority opinion*: Critical	51%	39%	6%	3%	2%	*Panel majority opinion*: Critical
	Academics	64%	32%	5%	0%	0%	*Groups majority opinions:* All Critical, except Buyers and policymakers (=50%).	57%	38%	0%	0%	5%	*Groups majority opinions:* All Critical, except Industry and Patients and citizens.
	Buyers and policymakers	50%	25%	13%	0%	13%		67%	17%	17%	0%	0%	
	Industry	67%	25%	8%	0%	0%		36%	45%	18%	0%	0%	
	Patients and citizens	73%	15%	6%	0%	6%		48%	36%	6%	6%	3%	
	Healthcare professionals	62%	38%	0%	0%	0%		52%	42%	4%	2%	0%	
Adverse events for the patient	Overall panel	81%	17%	1%	0%	2%	*Panel majority opinion*: Critical	66%	25%	3%	3%	3%	*Panel majority opinion*: Critical
	Academics	86%	14%	0%	0%	0%	*Groups majority opinions:* All Critical	67%	29%	0%	0%	5%	*Groups majority opinions:* All Critical, except Buyers and policymakers (=50%) and Industry.
	Buyers and policymakers	63%	38%	0%	0%	0%		50%	33%	17%	0%	0%	
	Industry	75%	25%	0%	0%	0%		45%	45%	0%	9%	0%	
	Patients and citizens	76%	15%	3%	0%	6%		61%	24%	3%	6%	6%	
	Healthcare professionals	87%	13%	0%	0%	0%		75%	19%	2%	2%	2%	
Quality of the available scientific evidence	Overall panel	34%	61%	3%	0%	2%	*Panel majority opinion:* Fundamental	45%	50%	4%	0%	2%	*Panel majority opinion:* No majority
	Academics	45%	55%	0%	0%	0%	*Groups majority opinions:* All Fundamental, except Industry (=50%).	62%	29%	5%	0%	5%	*Groups majority opinions:* Critical for Academics and Industry, Fundamental for Patients and citizens and Healthcare professionals. No majority for Buyers and policymakers.
	Buyers and policymakers	25%	63%	13%	0%	0%		50%	33%	17%	0%	0%	
	Industry	42%	50%	8%	0%	0%		55%	36%	9%	0%	0%	
	Patients and citizens	15%	76%	3%	0%	6%		33%	58%	6%	0%	3%	
	Healthcare professionals	40%	58%	2%	0%	0%		42%	58%	0%	0%	0%	
Target population	Overall panel	8%	60%	28%	1%	3%	*Panel majority opinion:* Fundamental	16%	57%	23%	3%	1%	*Panel majority opinion:* Fundamental
	Academics	18%	68%	14%	0%	0%	*Groups majority opinions:* All Fundamental, except Buyers and policymakers (=50%).	24%	67%	5%	0%	5%	*Groups majority opinions:* All Fundamental, except Buyers and policymakers (=50%) and Industry.
	Buyers and policymakers	25%	50%	13%	0%	13%		17%	50%	0%	33%	0%	
	Industry	17%	58%	25%	0%	0%		36%	36%	18%	9%	0%	
	Patients and citizens	0%	52%	39%	0%	9%		6%	55%	36%	3%	0%	
	Healthcare professionals	4%	63%	31%	2%	0%		15%	60%	25%	0%	0%	
Impact of the disease	Overall panel	15%	61%	20%	2%	2%	*Panel majority opinion:* Fundamental	23%	61%	12%	4%	1%	*Panel majority opinion:* Fundamental
	Academics	23%	50%	23%	0%	5%	*Groups majority opinions:* All Fundamental, except Academics (=50%) and Industry.	38%	43%	14%	0%	5%	*Groups majority opinions:* Fundamental for Patients and citizens and Healthcare professionals. No majority for the remaining groups (=50% on Fundamental for Buyers and policymakers).
	Buyers and policymakers	13%	75%	0%	13%	0%		33%	50%	0%	17%	0%	
	Industry	50%	42%	8%	0%	0%		45%	36%	9%	9%	0%	
	Patients and citizens	6%	61%	27%	0%	6%		12%	67%	15%	6%	0%	
	Healthcare professionals	10%	67%	21%	2%	0%		17%	71%	10%	2%	0%	
Patient-reported outcomes	Overall panel	17%	57%	21%	1%	3%	*Panel majority opinion:* Fundamental	20%	45%	26%	8%	1%	*Panel majority opinion:* No majority
	Academics	27%	55%	18%	0%	0%	*Groups majority opinions:* All Fundamental, except Industry (=50%).	29%	48%	19%	0%	5%	*Groups majority opinions:* No majority for all
	Buyers and policymakers	13%	63%	13%	0%	13%		17%	33%	33%	17%	0%	
	Industry	42%	50%	8%	0%	0%		27%	27%	27%	18%	0%	
	Patients and citizens	9%	61%	18%	3%	9%		15%	42%	33%	9%	0%	
	Healthcare professionals	13%	58%	29%	0%	0%		19%	50%	23%	8%	0%	
Quality of life for the patient	Overall panel	33%	58%	6%	0%	2%	*Panel majority opinion:* Fundamental	34%	50%	8%	7%	1%	*Panel majority opinion:* No majority
	Academics	55%	41%	5%	0%	0%		52%	38%	5%	0%	5%	
	Buyers and policymakers	50%	25%	13%	0%	13%	*Groups majority opinions:* All Fundamental, except Academics (Critical) and Buyers and policymakers.	33%	50%	0%	17%	0%	*Groups majority opinions:* Critical for Academics, Fundamental for Patients and citizens and Healthcare professionals. No majority for the remaining groups.
	Industry	33%	67%	0%	0%	0%		27%	36%	18%	18%	0%	
	Patients and citizens	27%	64%	3%	0%	6%		27%	58%	12%	3%	0%	
	Healthcare professionals	25%	65%	10%	0%	0%		33%	52%	6%	8%	0%	
Space for innovation for the healthcare organisation	Overall panel	3%	38%	50%	6%	4%	*Panel majority opinion:* No majority	7%	46%	40%	5%	2%	*Panel majority opinion:* No majority
	Academics	0%	27%	59%	9%	5%	*Groups majority opinions:* Complementary for Academics, Buyers and policymakers and Patients and citizens. No majority for the remaining groups.	0%	48%	48%	0%	5%	*Groups majority opinions:* Fundamental for Healthcare professionals, Complementary for Buyers and policymakers. No majority for the remaining groups.
	Buyers and policymakers	0%	25%	75%	0%	0%		0%	17%	83%	0%	0%	
	Industry	8%	42%	33%	17%	0%		18%	45%	27%	9%	0%	
	Patients and citizens	0%	33%	55%	0%	12%		12%	39%	42%	3%	3%	
	Healthcare professionals	6%	46%	42%	6%	0%		4%	54%	33%	8%	0%	
Clinical guidelines	Overall panel	13%	50%	31%	2%	2%	*Panel majority opinion:* No majority	16%	61%	19%	3%	1%	*Panel majority opinion:* Fundamental
	Academics	32%	18%	41%	5%	5%	*Groups majority opinions:* Fundamental for Buyers and policymakers, Industry and Patients and citizens. No majority for the remaining groups.	38%	38%	19%	0%	5%	*Groups majority opinions:* All Fundamental, except Academics.
	Buyers and policymakers	13%	63%	25%	0%	0%		0%	83%	17%	0%	0%	
	Industry	17%	67%	8%	8%	0%		9%	73%	9%	9%	0%	
	Patients and citizens	3%	64%	27%	0%	6%		12%	64%	21%	3%	0%	
	Healthcare professionals	12%	50%	37%	2%	0%		13%	65%	21%	2%	0%	
Financing	Overall panel	11%	49%	33%	3%	4%	*Panel majority opinion:* No majority	16%	50%	28%	3%	3%	*Panel majority opinion:* No majority
	Academics	9%	36%	50%	0%	5%	*Groups majority opinions:* Fundamental for Industry and Patients and citizens. No majority for the remaining groups.	29%	33%	29%	0%	10%	*Groups majority opinions:* All Fundamental, except Academics and Industry.
	Buyers and policymakers	25%	38%	25%	0%	13%		33%	67%	0%	0%	0%	
	Industry	17%	67%	17%	0%	0%		18%	45%	18%	18%	0%	
	Patients and citizens	6%	58%	24%	6%	6%		12%	58%	27%	3%	0%	
	Healthcare professionals	12%	46%	37%	4%	2%		10%	52%	33%	2%	2%	

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 8 of 15

Table 2 (continued)

Criterion	Group						Majority opinion						Majority opinion
Public health interest	Overall panel	14%	58%	25%	1%	2%	*Panel majority opinion:* Fundamental	24%	58%	17%	1%	1%	*Panel majority opinion:* Fundamental
	Academics	14%	59%	27%	0%	0%	*Groups majority opinions:* All Fundamental, except Industry.	29%	48%	19%	0%	5%	*Groups majority opinions:* Critical for Industry, Fundamental for Patients and citizens and Healthcare professionals. No majority for the remaining groups (=50% on Fundamental for Buyers and policymakers).
	Buyers and policymakers	0%	63%	38%	0%	0%		33%	50%	17%	0%	0%	
	Industry	50%	25%	25%	0%	0%		55%	27%	18%	0%	0%	
	Patients and citizens	15%	67%	12%	0%	6%		15%	73%	12%	0%	0%	
	Healthcare professionals	8%	60%	31%	2%	0%		19%	60%	19%	2%	0%	
Budget impact to the health system	Overall panel	9%	56%	32%	2%	2%	*Panel majority opinion:* Fundamental	19%	45%	32%	3%	1%	*Panel majority opinion:* No majority
	Academics	14%	59%	27%	0%	0%	*Groups majority opinions:* All Fundamental, except Industry.	24%	57%	14%	0%	5%	*Groups majority opinions:* Fundamental for Academics and Buyers and policymakers. No majority for the remaining groups.
	Buyers and policymakers	13%	63%	25%	0%	0%		33%	67%	0%	0%	0%	
	Industry	25%	42%	33%	0%	0%		27%	45%	27%	0%	0%	
	Patients and citizens	3%	61%	24%	6%	6%		9%	48%	36%	6%	0%	
	Healthcare professionals	6%	54%	40%	0%	0%		21%	35%	42%	2%	0%	
Equity	Overall panel	13%	59%	21%	3%	4%	*Panel majority opinion:* Fundamental	16%	55%	20%	6%	3%	*Panel majority opinion:* Fundamental
	Academics	9%	59%	27%	0%	5%	*Groups majority opinions:* All Fundamental	14%	52%	24%	0%	10%	*Groups majority opinions:* All Fundamental
	Buyers and policymakers	13%	75%	0%	0%	13%		17%	83%	0%	0%	0%	
	Industry	17%	58%	17%	8%	0%		18%	55%	9%	9%	9%	
	Patients and citizens	21%	52%	21%	0%	6%		12%	58%	24%	6%	0%	
	Healthcare professionals	8%	62%	23%	6%	2%		19%	52%	21%	8%	0%	
Market competitiveness	Overall panel	5%	27%	46%	20%	2%	*Panel majority opinion:* No majority	7%	31%	45%	16%	1%	*Panel majority opinion:* No majority
	Academics	0%	32%	32%	32%	5%	*Groups majority opinions:* Fundamental for Industry, Complementary for Healthcare Professionals. No majority for the remaining groups.	0%	33%	38%	24%	5%	*Groups majority opinions:* Complementary for Patients and citizens. No majority for the remaining groups.
	Buyers and policymakers	25%	13%	38%	13%	13%		17%	33%	17%	33%	0%	
	Industry	8%	58%	25%	8%	0%		9%	36%	36%	18%	0%	
	Patients and citizens	3%	18%	48%	27%	3%		6%	24%	55%	15%	0%	
	Healthcare professionals	4%	25%	56%	15%	0%		8%	33%	48%	10%	0%	
Medical or technical complications for the patient	Overall panel	68%	28%	2%	0%	2%	*Panel majority opinion:* Critical	58%	33%	4%	3%	2%	*Panel majority opinion:* Critical
	Academics	55%	41%	0%	0%	5%	*Groups majority opinions:* All Critical, except for Buyers and policymakers (=50%).	62%	24%	10%	0%	5%	*Groups majority opinions:* All Critical, except for Buyers and policymakers (Fundamental).
	Buyers and policymakers	50%	38%	13%	0%	0%		33%	67%	0%	0%	0%	
	Industry	67%	33%	0%	0%	0%		55%	27%	0%	18%	0%	
	Patients and citizens	73%	21%	0%	0%	6%		55%	36%	3%	3%	3%	
	Healthcare professionals	73%	25%	2%	0%	0%		63%	31%	4%	2%	0%	
Stakeholders agreement on the adoption of the medical device	Overall panel	9%	56%	27%	6%	3%	*Panel majority opinion:* Fundamental	12%	59%	24%	3%	2%	*Panel majority opinion:* Fundamental
	Academics	14%	50%	23%	9%	5%	*Groups majority opinions:* Fundamental for Patients and citizens and Healthcare professionals. No majority for the remaining groups (=50% Fundamental for Academics and Industry).	14%	52%	24%	5%	5%	*Groups majority opinions:* All Fundamental, except Industry.
	Buyers and policymakers	0%	38%	50%	0%	13%		0%	67%	33%	0%	0%	
	Industry	8%	50%	42%	0%	0%		27%	45%	27%	0%	0%	
	Patients and citizens	9%	61%	15%	9%	6%		9%	64%	21%	3%	3%	
	Healthcare professionals	8%	60%	29%	4%	0%		10%	60%	25%	4%	0%	
Environmental impact of the production and use of the medical device	Overall panel	3%	33%	52%	9%	2%	*Panel majority opinion:* Complementary	9%	39%	45%	7%	1%	*Panel majority opinion:* No majority
	Academics	5%	41%	50%	5%	0%	*Groups majority opinions:* Complementary for Industry and Healthcare Professionals. No majority for the remaining groups (=50% Complementary for Academics and Buyers and policymakers).	5%	38%	48%	5%	5%	*Groups majority opinions:* No majority for all.
	Buyers and policymakers	13%	0%	50%	38%	0%		33%	17%	50%	0%	0%	
	Industry	0%	25%	67%	8%	0%		9%	45%	36%	9%	0%	
	Patients and citizens	0%	36%	48%	9%	6%		6%	42%	42%	9%	0%	
	Healthcare professionals	4%	35%	52%	8%	2%		10%	38%	46%	6%	0%	
Efficiency	Overall panel	36%	53%	9%	1%	2%	*Panel majority opinion:* Fundamental	57%	37%	5%	0%	1%	*Panel majority opinion:* Critical
	Academics	50%	50%	0%	0%	0%	*Groups majority opinions:* Fundamental for Buyers and policymakers and Patients and citizens. No majority for the remaining groups (=50% Fundamental for Academics).	62%	33%	0%	0%	5%	*Groups majority opinions:* All Critical, except Buyers and policymakers (=50%).
	Buyers and policymakers	38%	63%	0%	0%	0%		50%	33%	17%	0%	0%	
	Industry	50%	42%	8%	0%	0%		82%	18%	0%	0%	0%	
	Patients and citizens	18%	64%	12%	0%	6%		52%	42%	6%	0%	0%	
	Healthcare professionals	38%	48%	12%	2%	0%		54%	40%	6%	0%	0%	
Capacity of the health system	Overall panel	13%	60%	20%	4%	3%	*Panel majority opinion:* Fundamental	14%	63%	20%	2%	1%	*Panel majority opinion:* Fundamental
	Academics	18%	59%	14%	5%	5%	*Groups majority opinions:* All Fundamental, except Buyers and policymakers (=50%) and Industry.	14%	62%	14%	5%	5%	*Groups majority opinions:* All Fundamental, except Industry.
	Buyers and policymakers	25%	50%	13%	0%	13%		17%	67%	17%	0%	0%	
	Industry	33%	42%	25%	0%	0%		36%	36%	27%	0%	0%	
	Patients and citizens	9%	64%	12%	9%	6%		15%	67%	15%	3%	0%	
	Healthcare professionals	8%	63%	27%	2%	0%		8%	67%	25%	0%	0%	
Cost of the medical device (including complementary equipment)	Overall panel	19%	60%	17%	1%	3%	*Panel majority opinion:* Fundamental	25%	55%	16%	3%	2%	*Panel majority opinion:* Fundamental
	Academics	36%	55%	5%	0%	5%	*Groups majority opinions:* All Fundamental	48%	48%	0%	0%	5%	*Groups majority opinions:* All Fundamental, except Buyers and policymakers (Critical) and Academics.
	Buyers and policymakers	25%	75%	0%	0%	0%		67%	33%	0%	0%	0%	
	Industry	17%	75%	8%	0%	0%		18%	64%	9%	9%	0%	
	Patients and citizens	9%	55%	27%	0%	9%		18%	58%	18%	3%	3%	
	Healthcare professionals	17%	60%	21%	2%	0%		17%	56%	25%	2%	0%	
Cost of procedure without the cost of the medical device	Overall panel	13%	57%	24%	3%	4%	*Panel majority opinion:* Fundamental	21%	55%	17%	6%	2%	*Panel majority opinion:* Fundamental
	Academics	23%	68%	0%	5%	5%	*Groups majority opinions:* All Fundamental, except Patients and citizens.	33%	48%	10%	5%	5%	*Groups majority opinions:* All Fundamental, except Academics and Buyers and policymakers (= 50%).
	Buyers and policymakers	13%	88%	0%	0%	0%		50%	50%	0%	0%	0%	
	Industry	17%	67%	17%	0%	0%		18%	55%	18%	9%	0%	
	Patients and citizens	12%	42%	36%	0%	9%		21%	52%	21%	3%	3%	
	Healthcare professionals	8%	54%	31%	6%	2%		13%	60%	19%	8%	0%	

'Don't know…' column present the answers for 'Don't know / Don't want to answer'. The relevance levels gathering majority are highlighted in dark grey and, if there is no majority, the relevance levels gathering most answers are highlighted in light grey

gathering opinion among 'Critical' and 'Fundamental' and Healthcare professionals among 'Fundamental' and 'Complementary'. For 'In vitro tests based on biomarkers', a statistically significant difference was found between Academics and Patients and citizens in 'Target population' ($p = 0.0370$), with both groups gathering majority in 'Fundamental' but Academics gathering opinion among 'Critical' and 'Fundamental' and Patients and citizens among 'Fundamental' and 'Complementary', and between Academics and Healthcare professionals in 'Cost of the medical device (including complementary equipment)' ($p = 0.00718$), with Academics' answers gathering no majority and Healthcare professionals gathering majority in 'Fundamental'. For the remaining aspects, the differences across groups were not statistically significant.

Agreement within groups of stakeholders for each of the two types of medical devices

All groups presented a higher Gwet's AC2 coefficient for 'implantable' than for 'in vitro' devices (see Additional file 3). Irrespectively of the type of medical device, the agreement increased in round 2, suggesting answers' convergence. According to both benchmarking methods, Landis and Koch [44] and Gwet's proposed benchmarking method [41], the calculated Gwet's coefficients describe a strength of agreement from 'moderate' to 'substantial' among all stakeholders' groups. The agreement level assessed first, following Landis and Koch, and after, using Gwet's proposed benchmarking method, comprised the following differences: for 'Implantable medical devices', the agreement changed from 'substantial' to 'moderate' in the Buyers and policymakers group (in round 1) and for 'In vitro tests based on biomarkers', the agreement changed from 'substantial' to 'moderate' in the Buyers and policymakers group (in round2). Both benchmarking methods resulted in a 'moderate' agreement in the Industry group (in both rounds for 'in vitro' devices) and in the Buyers and policymakers group (in round 1, for 'in vitro' devices). A 'substantial' agreement was observed in all the remaining cases.

Common relevance for both types of medical devices

Similar results were retrieved for both types of medical devices: five aspects were 'Critical' and 11 'Fundamental', according to the panel – in three 'Critical' and in one 'Fundamental', the majority was observed in all stakeholders' groups. In the remaining 18 aspects, seven aspects did not gather majority in any type of device, 10 gathered majority for only one type and 1 gathered majority on a different relevance level across both types of devices. These results, previously detailed in Table 2,

are synthesised in Table 3 together with the common relevance level.

Discussion

Based upon the opinions of the distinct stakeholders' groups involved in the Web-Delphi process, this study, firstly, explored the relevance of 34 aspects for the evaluation of each of two types of medical devices, according to the panel and per group of stakeholders; secondly, analysed the level of agreement within stakeholders' groups; and, finally, concluded about which aspects gathered a common relevance across the two types of medical devices. This work was developed to form a basis for discussing HTA processes and the construction of value models to evaluate medical devices. It is hereinafter discussed in terms of stakeholders' views on the aspects to consider in medical devices evaluation, implications for policy, and limitations of the study.

Stakeholders' views on the aspects to consider in medical devices evaluation

The Delphi panel was composed of distinct stakeholders' groups. Analysing the distribution of answers of the panel and per groups, results show one of four situations: (1) there is a panel majority opinion with all groups presenting majority on the same relevance level, (2) there is a panel majority opinion but not all groups present majority on the same relevance level, (3) there is no panel majority opinion but some groups present majority on a relevance level, and (4) there is no panel majority opinion nor majority in any group. Considering the meaning of the relevance levels (presented in *Overview of the Web-Delphi process*, in the *Methods* section), these results suggest that participants consider there are aspects that must always be part of the basis of evaluation: that is the case of aspects assigned 'Fundamental' or 'Critical' by the panel, for instance, 'User-friendliness for the healthcare professional' for 'implantable' devices and 'Time between procedure and results' for 'in vitro' devices. Furthermore, participants consider some of these aspects can even preclude the evaluation if there is no data for assessing them – this applies to aspects set as 'Critical' –, for instance, 'Specific features of the medical device' for 'implantable' devices and 'Sensitivity and Specificity' for 'in vitro' devices. Additionally, participants' answers suggest that there are 'Complementary' aspects, i.e., aspects that can add some value but will not always be part of the basis of evaluation, for instance 'Environmental impact of the production and use of the medical device' and 'Learning curve of the healthcare professional' for 'implantable' and 'in vitro' devices, respectively. This can be seen in Table 3 that presents the panel majority opinion and systematizes

Table 3 Summary of conclusions about the common relevance of aspects for both types of medical devices

Common relevance	Aspect	Implantable medical devices (127 participants)							In vitro diagnostic tests based on biomarkers (119 participants)						
		Panel majority opinion	Groups majority opinions (✓ if aligned with panel majority opinion)						Panel majority opinion	Groups majority opinions (if aligned with panel majority opinion)					
			A	BP	I	PC	HP			A	BP	I	PC	HP	
Critical	Technical performance of the medical device	Critical	✓	✓	✓	✓	✓		Critical	✓	✓	✓	✓	✓	
	Clinical efficacy and/or effectiveness	Critical	✓	✓	✓	✓	✓		Critical	✓	✓	✓	✓	✓	
	Risk analysis	Critical	✓	✓	✓	✓	✓		Critical	✓	✓	✓	✓	✓	
	Adverse events for the patient	Critical	✓	✓	✓	✓	✓		Critical	✓	✓	✓	✓	✓	
	Medical or technical complications for the patient	Critical	✓	✓	✓	✓	✓		Critical	✓	Fu	✓	✓	✓	
Fundamental	User-friendliness for the healthcare professional	Fundamental	✓	✓	✓	✓	✓		Fundamental	✓		✓	✓	✓	
	Need for training of the healthcare professional	Fundamental	✓			✓	✓		Fundamental	✓			✓	✓	
	Comfort for the patient	Fundamental	✓		✓	✓	✓		Fundamental	✓			✓	✓	
	Target population	Fundamental		✓	✓	✓	✓		Fundamental	✓			✓	✓	
	Impact of the disease	Fundamental	✓			✓	✓		Fundamental			✓		✓	
	Public health interest	Fundamental	✓	✓		✓	✓		Fundamental			Cr	✓	✓	
	Equity	Fundamental	✓	✓	✓	✓	✓		Fundamental	✓	✓	✓	✓	✓	
	Stakeholders agreement on the adoption of the medical device	Fundamental			✓	✓	✓		Fundamental	✓	✓		✓	✓	
	Capacity of the health system	Fundamental	✓			✓	✓		Fundamental	✓	✓		✓	✓	
	Cost of the medical device (including complementary equipment)	Fundamental	✓	✓	✓	✓	✓		Fundamental		Cr	✓	✓	✓	
	Cost of procedure without the cost of the medical device	Fundamental	✓	✓	✓	✓	✓		Fundamental			✓	✓	✓	

Table 3 (continued)

| Common relevance | Aspect | Implantable medical devices (127 participants) | | | | | | | In vitro diagnostic tests based on biomarkers (119 participants) | | | | | | |
| | | Panel majority opinion | Groups majority opinions (✓ if aligned with panel majority opinion) | | | | | | Panel majority opinion | Groups majority opinions (if aligned with panel majority opinion) | | | | | |
			A	BP	I	PC	HP			A	BP	I	PC	HP
-	Specific features of the medical device	Critical	✓	✓	✓	✓	✓		No majority		Cr		Fu	
	Regulatory status of the medical device	No majority		Fu	Cr	Fu			No majority		Cr			
	Sensitivity and Specificity	No majority	Cr	Fu	Cr				Critical	✓	✓	✓	✓	✓
	Time between procedure and results	No majority					Fu		Fundamental	✓	✓	✓	✓	✓
	Learning curve of the healthcare professional	No majority		Co		Co	Fu		Complementary	✓	✓	✓	✓	
	Exposure of the healthcare professional to physical or chemical agents	Fundamental	✓		✓	✓			No majority	Fu			Fu	
	Workload for the healthcare professional	No majority					Fu		No majority	Fu	Co		Fu	
	Connectivity	No majority			Fu	Fu	Co		No majority	Co				
	Quality of the available scientific evidence	Fundamental	✓	✓		✓	✓		No majority	Cr		Cr	Fu	Fu
	Patient-reported outcomes	Fundamental	✓	✓	✓	✓	✓		No majority					
	Quality of life for the patient	Fundamental	Cr		✓	✓	✓		No majority	Cr			Fu	Fu
	Space for innovation for the healthcare organization	No majority	Co	Co		Co			No majority		Co			Fu
	Clinical guidelines	No majority		Fu	Fu	Fu			Fundamental		✓	✓	✓	✓
	Financing	No majority		Fu	Fu	Fu			No majority		Fu		Fu	Fu
	Budget impact to the health system	Fundamental	✓	✓		✓	✓		No majority	Fu	Fu			
	Market competitiveness	No majority			Fu		Co		No majority				Co	
	Environmental impact of the production and use of the medical device	Complementary			✓		✓		No majority					
	Efficiency	Fundamental		✓		✓			Critical	✓		✓	✓	✓

A Academics, *BP* Buyers and policymakers, *I* Industry, *PC* Patients and citizens, *HP* Healthcare professionals, *Cr* Critical, *Fu* Fundamental, *Co* Complementary

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 12 of 15

the groups majorities aligned with the panel. In general, stakeholders' groups did not present obvious contradictory opinions, as even the aspects with no panel majority opinion gathered most groups answers around the same two consecutive relevance levels (as presented in Table 2). The Kruskal–Wallis test followed by Dunn's-Bonferroni post hoc method allowed to identify only eight aspects (six in 'implantable' and two in 'in vitro' devices) with statistically significant differences across groups, but these differences were always observed across only one pair of groups, and the inter-rater reliability calculated with Gwet's AC2 agreement coefficient showed a strength of agreement from moderate to substantial within each group, suggesting an alignment of the panel and within groups. Despite the general alignment of the groups, the reasons underlying the observed differences may be of interest for further research [46].

This Web-Delphi process collected opinions not only from different stakeholders' groups but also for two types of devices, a therapeutic and a diagnostic type of device. Comparing the opinions across both types of devices, results show that the panel attributed a common relevance level for 16 aspects, five 'Critical' (two agreed by all stakeholders' groups) and 11 'Fundamental' (one agreed by all groups) (see Table 3). Examples of this are the 'Clinical efficacy and/or effectiveness' considered 'Critical' (majority in all groups), or the 'Comfort for the patient' considered 'Fundamental' (not getting majority in all groups but gathering a panel majority opinion). The former aspect is aligned with economic evaluation literature [6] centred into the effectiveness and costs of technologies whereas the latter is not explicitly considered in such methods. Moreover, many other aspects were recognised as relevant by the participants of our study, suggesting the need to formally consider a larger number of aspects in the evaluation of medical devices. This need has been recognised in literature [3, 47], by authors advocating for the use of MCDA in HTA [2], such as the ISPOR (The Professional Society for Health Economics and Outcomes Research) Medical Devices and Diagnostics Special Interest Group [48], by authors developing value framework models for evaluating medical devices, such as in the HTA Core Model from EUnetHTA (European network for Health Technology Assessment) [49], and also by the review on value assessment frameworks of Zhang et al. [22] that covered 19 studies addressing health technologies in general and 38 addressing specific types of health technologies (mainly drugs). Four of the frameworks reported in that review targeted diagnostic or genetic tests and one targeted nondrug health technologies, with evaluation aspects included varying between three and 16 and covering different devices' features, namely, their medical benefit, the adverse effects,

the quality of life and satisfaction of the patient, and the costs. Our study, besides validating this need with a large and diverse panel of stakeholders, adds additional value aspects not identified in the existing frameworks' literature, e.g., regarding environmental impact and aspects related with devices' usage by the healthcare professional, such as user-friendliness, the learning curve, the training and the workload.

The list of aspects included in our study tries to be purposefully inclusive, which brings the advantage of being as complete as possible but the disadvantage of entailing potential overlap in some aspects. To evolve towards the construction of a multidimensional framework or of multicriteria models, the identified aspects would require further work and restructuring, eventually combining and clarifying aspects and exploring how to measure them in practice [32] (for e.g., understanding what participants have in mind when considering the sensitivity and specificity of an implantable medical device as relevant). Nonetheless, our work provides important insights to inform such a framework development. In 44 frameworks reviewed by Zhang et al. [22], value aspects were identified through literature review, engagement of stakeholders, or a combination of both, but only four frameworks involved patients or citizens in aspects' identification. Our study explored a way to collect the wide range and diversity of stakeholders' perspectives, including patients and citizens, adding to the discussion on how to bring these insights into HTA for standardising and bringing guidance and transparency to the evaluation of medical devices [19, 28, 50] and how to include stakeholders' views to inform HTA and adoption decisions [23, 26].

Methodologically, through the Web-Delphi it was thus possible, first, to involve a large and heterogeneous group of HTA stakeholders, enabling them to interact and learn with each other by sharing their views and build an agreement about the relevance of most aspects. Second, to draw conclusions about differences in opinion between stakeholders and across types of devices. Third, it has shown in which aspects there is a panel majority opinion. All of this provides input information for additional research on how to develop multidimensional evaluation models and frameworks, and assists in planning future directions of research.

Implications for policy

This study shows that it is possible to gather the views of distinct stakeholders' groups in a structured format, producing results that can be more widely used within HTA processes, as deemed as relevant by several authors [24–27, 51]. All aspects were considered to some extent relevant, and some aspects gathered the same relevance

level irrespectively of the type of medical device under analysis. Accordingly, approaches to assess medical devices value need to consider a broader range of aspects and the specificities of distinct types of devices. Despite the heterogeneity of this type of technologies [12], there seems to be possible to attain some systematization and common standards, so asked in literature [4, 9, 20, 52]. Nevertheless, one should recall that the evaluation may be affected by the context [53], and that 'Implantable medical devices' and 'In vitro tests based on biomarkers' still comprise diverse devices, which needs to be considered when interpreting results.

Limitations of the study

Several limitations should be acknowledged in this study. Firstly, this study takes place in Portugal, having only national participants, and thus results can be context- and/or country-dependent. Nevertheless, the list of aspects was based on international peer-reviewed literature which brings useful information to inform the discussion on HTA for medical devices, beyond the considered country and context. Secondly, as the Delphi process is highly dependent on the availability and commitment of participants [54], there was not a balance of participants across stakeholder groups, with the Buyers and policymakers and the Industry having a lower representation. This unbalance is somehow usual in Delphi processes as panels are purposive or convenience samples, not aiming to be representative samples of populations [54]. Furthermore, Delphi literature does not present unequivocal recommendations for the sample size, with studies suggesting ranges from five to more than one thousand participants [54, 55]. To try to mitigate the unbalance as much as possible, invitations and reminders for participation were sent. Thirdly, and still related with the Delphi process, it is important to acknowledge shortcomings of the method, namely the possibility of occurring cognitive biases and other behavioural influences during the process (such as egocentric discounting or the influence of majority positions) due to the freely online interactions among participants, which can also lead to answers not completely clarified by participants [56], or the possibility of information overload due to the high number of aspects to be analysed by the participants, which could become tiresome and cognitively challenging for them [57]. To avoid the occurrence of such shortcomings, not only the panel of participants was heterogeneous but also the aspects were organised, during the validation with experts, so that it would be easier, to the best of their knowledge, to follow the exercise. Additionally, participants could also rate each type of medical device in different time periods by

re-accessing the platform, or even answering only one type of medical device if they felt more comfortable. Finally, the initial list of aspects could be biased as it was the result of a literature search followed by the validation by experts of the HTA agency. To overcome this possibility, participants could suggest additional aspects during the process, through the comments option, which was not observed.

Conclusion

One hundred thirty-four participants, belonging to different groups of health stakeholders, recognised many aspects, besides costs and effectiveness, as relevant for the evaluation of 'implantable' and 'in vitro' medical devices. Results suggest the need to formally consider a larger number of aspects in such evaluations. The results from this study have implications for the development of multidimensional value frameworks and models in HTA for medical devices, contributing to guide evidence collection to inform evaluators. In future research, these results could be discussed by HTA agencies and decision-makers, namely for understanding the extent to which these findings can be applied for other types of medical devices and embedded within HTA processes. Additionally, research has been conducted within the scope of the MEDI-VALUE project and confirmed that the results of this study are a useful starting point for the development of multicriteria models to evaluate specific medical devices in real contexts, but this should be extended to other real-world settings.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12913-023-09550-0.

> **Additional file 1: Table 1.1.** List and descriptions of the 34 aspects used in the Web-Delphi process, translated to English (original list and descriptions delivered in Portuguese).
>
> **Additional file 2: Table 2.1.** Statistical tests results for each type of medical devices: Kruskal-Wallis with 4 degrees of freedom and Dunn post hoc test, corrected with Bonferroni test.
>
> **Additional file 3: Table 3.1.** Percent agreement, Gwet's coefficient of groups of stakeholders for both types of medical devices in both rounds of the Web-Delphi process, and strength of agreement according to Landis and Koch and to Gwet's proposed benchmarking method.

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 14 of 15

Availability of data and materials
All data generated or analysed during this study are included within the article and its additional files.

Declarations

Ethics approval and consent to participate
All methods were carried out in accordance with relevant guidelines and regulations. This study was approved by the Ethics Committee of Instituto Superior Técnico, University of Lisbon (reference no. 14/2019 (CE-IST)). Informed consent was obtained from all individual participants in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal. [2]iBB- Institute for Bioengineering and Biosciences and i4HB- Associate Laboratory Institute for Health and Bioeconomy, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1, Lisbon 1049-001, Portugal. [3]Infarmed, I. P., Lisbon, Portugal. [4]LSE Health-Medical Technology Research Group (MTRG), London School of Economics, Houghton St, London WC2A 2AE, UK.

References

1. Taylor RS, Iglesias CP. Assessing the clinical and cost-effectiveness of medical devices and drugs: are they that different? Value Health. 2009;12(4):404–6.
2. Oliveira MD, Mataloto I, Kanavos P. Multi-criteria decision analysis for health technology assessment: addressing methodological challenges to improve the state of the art. Eur J Health Econ. 2019;20:891–918.
3. Fuchs S, Olberg B, Panteli D, Perleth M, Busse R. HTA of medical devices: Challenges and ideas for the future from a European perspective. Health Policy. 2017;121(3):215–29.
4. Blüher M, Saunders SJ, Mittard V, Torrejon Torres R, Davis JA, Saunders R. Critical Review of European Health-Economic Guidelines for the Health Technology Assessment of Medical Devices. Front Med. 2019;6:278.
5. Rotter JS, Foerster D, Bridges JF. The changing role of economic evaluation in valuing medical technologies. Expert Rev Pharmacoecon Outcomes Res. 2012;12(6):711–23.
6. 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.
7. Ciani O, Wilcher B, Blankart CR, Hatz M, Rupel VP, Erker RS, Varabyova Y, Taylor RS. Health technology assessment of medical devices: a survey of non-European union agencies. Int J Technol Assess Health Care. 2015;31(3):154–65.
8. Tarricone R, Ciani O, Torbica A, Brouwer W, Chaloutsos G, Drummond MF, Martelli N, Persson U, Leidl R, Levin L, et al. Lifecycle evidence requirements for high-risk implantable medical devices: a European perspective. Expert Rev Med Devices. 2020;17(10):993–1006.
9. Tarricone R, Torbica A, Drummond M. Key Recommendations from the MedtecHTA Project. Health Econ. 2017;26(S1):145–52.
10. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs–same or different? Value Health. 2009;12(4):402–4.
11. World Health Organization. Health technology assessment of medical devices. Geneva: World Health Organization; 2011.
12. WHO. Medical devices. [https://www.who.int/health-topics/medical-devices#tab=tab_1]. Accessed 08 June 2022.
13. Regulation (EU) 2017/745 of the European parliament and of the council of 5 April 2017 on medical devices, amending directive 2001/83/EC, regulation (EC) no 178/2002 and regulation (EC) No 1223/2009 and repealing council directives 90/385/EEC and 93/42/EEC [2017] OJ L117/1.
14. Regulation (EU) 2017/746 of the European parliament and of the council of 5 April 2017 on in vitro diagnostic medical devices and repealing directive 98/79/EC and commission decision 2010/227/EU [2017] OJ L117/176.
15. Regulation (EU) 2021/2282 of the European parliament and of the council of 15 december 2021 on health technology assessment and amending directive 2011/24/EU [2021] OJ L458/1.
16. Schnell-Inderst P, Mayer J, Lauterberg J, Hunger T, Arvandi M, Conrads-Frank A, Nachtnebel A, Wild C, Siebert U. Health technology assessment of medical devices: What is different? An overview of three European projects. Z Evid Fortbild Qual Gesundhwes. 2015;109(4):309–18.
17. Gomes M, Murray E, Raftery J. Economic Evaluation of Digital Health Interventions: Methodological Issues and Recommendations for Practice. Pharmacoeconomics. 2022;40(4):367–78.
18. Ciani O, Wilcher B, van Giessen A, Taylor RS. Linking the Regulatory and Reimbursement Processes for Medical Devices: The Need for Integrated Assessments. Health Econ. 2017;26(Suppl 1):13–29.
19. Baltussen R, Jansen MPM, Bijlmakers L, Grutters J, Kluytmans A, Reuzel RP, Tummers M, van der Wilt GJ. Value assessment frameworks for HTA agencies: The organization of evidence-informed deliberative processes. Value Health. 2017;20(2):256–60.
20. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, Rabbe S, Shatrov K, Zemplenyi A, Taylor RS. Surrogate endpoints in health technology assessment: an international review of methodological guidelines. Pharmacoeconomics. 2020;38(10):1055–70.
21. Tarricone R, Amatucci F, Armeni P, Banks H, Borsoi L, Callea G, Ciani O, Costa F, Federici C, Torbica A, et al. Establishing a national HTA program for medical devices in Italy: Overhauling a fragmented system to ensure value and equal access to new medical technologies. Health Policy. 2021;125(5):602–8.
22. Zhang M, Bao Y, Lang Y, Fu S, Kimber M, Levine M, Xie F. What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. Value Health. 2022;25(2):302–17.
23. van Voorn GA, Vemer P, Hamerlijnck D, Ramos IC, Teunissen GJ, Al M, Feenstra TL. The missing stakeholder group: why patients should be involved in health economic modelling. Appl Health Econ Health Policy. 2016;14(2):129–33.
24. Angelis A, Kanavos P. Value-based assessment of new medical technologies: towards a robust methodological framework for the application of multiple criteria decision analysis in the context of health technology assessment. Pharmacoeconomics. 2016;34(5):435–46.
25. O'Donnell JC, Pham SV, Pashos CL, Miller DW, Smith MD. Health technology assessment: lessons learned from around the world - an overview. Value Health. 2009;12:S1–5.
26. Mueller D, Tivey D, Croce D. Health-technology assessment: Its role in strengthening health systems in developing countries. South Afr J Public Health. 2017;2(1):6–11.
27. Mueller D, Pattinson RC, Hlongwane TM, Busse R, Panteli D. Portable continuous wave Doppler ultrasound for primary healthcare in South Africa: can the EUnetHTA Core Model guide evaluation before technology adoption? Cost Eff Resour Alloc. 2021;19:8.
28. Domingo C, Fernandez M, Garin N, Milara J, Moran I, Muerza I, Pacheco A, Teruel C, Bentley R, Subiran R et al. Determining what represents value

Freitas *et al. BMC Health Services Research*   (2023) 23:593

Page 15 of 15

in the treatment of refractory or unexplained chronic cough from the perspective of key stakeholders in Spain using multi-criteria decision analysis. Appl Health Econ Health Policy. 2023;21(1):119–310.

29. Ruseckaite R, Maharaj AD, Dean J, Krysinska K, Ackerman IN, Brennan AL, Busija L, Carter H, Earnest A. Forrest CB *et al*: Preliminary development of recommendations for the inclusion of patient-reported outcome measures in clinical quality registries. BMC Health Serv Res. 2022;22(1):276.

30. Bell E, Neri M, Steuten L. Towards a broader assessment of value in vaccines: the BRAVE way forward. Appl Health Econ Health Policy. 2022;20(1):105–17.

31. Bana e Costa CA, Oliveira MD, Vieira ACL, Freitas L, Rodrigues TC, Bana e Costa J, Freitas Â, Santana P. Collaborative development of composite indices from qualitative value judgements: The EURO-HEALTHY Population Health Index model. Eur J Oper Res. 2023;305(1):475–92.

32 Vieira ACL, Oliveira MD, Bana e Costa CA. Enhancing knowledge construction processes within multicriteria decision analysis: The collaborative value modelling framework. Omega. 2020;94:102047.

33. Linstone HA, Turoff M. Delphi: A brief look backward and forward. Technol Forecast Soc Change. 2011;78(9):1712–9.

34. MEDI-VALUE - Developing HTA tools to consensualise MEDIcal devices' VALUE through multicriteria decision analysis. [https://medivalue.tecnico.ulisboa.pt/]. Accessed 11 Feb 2020.

35. Linstone HA, Turoff A. The Delphi Method: Techniques and Applications. Addison-Wesley Publishing Company, Advanced Book Program; 2002.

36. Decision Eyes. Welphi. [https://www.welphi.com/en/Home.html]. Accessed 09 Mar 2022.

37. Bana e Costa CA, Corrêa EC, De Corte JM, Vansnick JC. Facilitating bid evaluation in public call for tenders: a socio-technical approach. Omega. 2002;30(3):227–42.

38. Marttunen M, Haag F, Belton V, Mustajoki J, Lienert J. Methods to inform the development of concise objectives hierarchies in multi-criteria decision analysis. Eur J Oper Res. 2019;277(2):604–20.

39. Oliveira MD, Vieira ACL, Dimitrovova K, Angelis A, Kanavos P, Bana e Costa CA. Multi-criteria evaluation framework - Advancing knowledge and MCDA tools to assist HTA agencies in evaluating medicines on a common basis, Report 7.2 of IMPACT HTA. 2021. https://doi.org/10.5281/zenodo.7471696.

40. Field A, Miles J, Field Z. Discovering Statistics Using R. SAGE Publications; 2012.

41. Gwet KL. Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters. 4th ed. USA: Advanced Analytics, LLC; 2014.

42 Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol. 2013;13:61.

43. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. 5th ed. Oxford: Oxford University Press; 2015.

44. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159–74.

45. Klein D. Implementing a general framework for assessing interrater agreement in Stata. Stand Genomic Sci. 2018;18(4):871–901.

46. Mott DJ, Ternent L, Vale L. Do preferences differ based on respondent experience of a health issue and its treatment? A case study using a public health intervention. Eur J Health Econ. 2023;24:413–23.

47. Trapero-Bertran M, Rodríguez-Martín B, López-Bastida J. What attributes should be included in a discrete choice experiment related to health technologies? A systematic literature review. PLoS ONE. 2019;14(7): e0219905.

48. Onwudiwe NC, Charter R, Gingles B, Abrishami P, Alder H, Bahkai A, Civic D, Kließ MK, Lessard C, Zema CL. Generating Appropriate and Reliable Evidence for Value Assessment of Medical Devices: An ISPOR Medical Devices and Diagnostics Special Interest Group Report. J Med Device. 2022;16(3):034701–10.

49. Kristensen FB, Lampe K, Wild C, Cerbo M, Goettsch W, Becla L. The HTA Core Model®—10 Years of Developing an International Framework to Share Multidimensional Value Assessment. Value Health. 2017;20(2):244–50.

50. Efthymiadou O, Mossman J, Kanavos P. Health related quality of life aspects not captured by EQ-5D-5L: Results from an international survey of patients. Health Policy. 2019;123(2):159–65.

51. Sorenson C, Drummond M, Kristensen FB, Busse R. How can the impact of health technology assessments be enhanced?, vol. EUR/07/5065810. Copenhagen: WHO Regional Office for Europe; 2008.

52. Drummond M, Tarricone R, Torbica A. European union regulation of health technology assessment: what is required for it to succeed? Eur J Health Econ. 2022;23(6):913–5.

53. Ghabri S, Josselin J-M, Le Maux B. Could or Should We Use MCDA in the French HTA Process? Pharmacoeconomics. 2019;37(12):1417–9.

54. Belton I, MacDonald A, Wright G, Hamlin I. Improving the practical application of the Delphi method in group-based judgment: A six-step prescription for a well-founded and defensible process. Technol Forecast Soc Change. 2019;147:72–82.

55 Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. BMC Med Res Methodol. 2005;5:37.

56. Makkonen M, Hujala T, Uusivuori J. Policy experts' propensity to change their opinion along Delphi rounds. Technol Forecast Soc Change. 2016;109:61–8.

57. Belton I, Wright G, Sissons A, Bolger F, Crawford MM, Hamlin I, Taylor Browne Lūka C, Vasilichi A. Delphi with feedback of rationales: How large can a Delphi group be such that participants are not overloaded, de-motivated, or disengaged? Technol Forecast Soc Change. 2021;170:120897.

Publisher's Note