Received: 5 March 2021 | Revised: 29 June 2021 | Accepted: 26 July 2021

DOI: 10.1002/mcda.1764

SPECIAL ISSUE ARTICLE

WILEY

Health Technology Assessment of medical devices. Overcoming the critical issues of current assessment

Salvatore Tallarico[1] | **Davide Aloini**[2] | **Riccardo Dulmin**[2] | **Simone Lazzini**[1] | **Valeria Mininno**[2] | **Luisa Pellegrini**[2]

[1]Department of Economics and Management, University of Pisa, Pisa, Italy

[2]Department of Energy, Systems, Territory and Constructions Engineering, University of Pisa, Pisa, Italy

Correspondence
Salvatore Tallarico, Department of Economics and Management, University of Pisa, Via Cosimo Ridolfi 10, 56124, Pisa, Italy.
Email: salvatore.tallarico@ec.unipi.it

Abstract

Medical devices are introduced in the market with an ever-increasing rate, with important benefits for the whole society's health. In the greatest majority of cases, although representing the entry point for devices, healthcare local institutions/hospitals do not receive any macro-level guidance from (inter)national authorities for their assessment. Hence, in the absence of a methodological support, local institutions/hospitals started to assess devices autonomously, building on Health Technology Assessment (HTA) tools based on Multi-Criteria Decision Analysis (MCDA), in order to take into-account the multifaceted aspects connected with devices. However, the strategies described so far in the scientific literature for implementing MCDA in local-based HTA of medical devices suffer from a harsh methodological weaknesses – that is, the use of bespoke criteria for the specific device to be assessed - that severely affect the evaluation of medical devices. Within this scenario, this work proposes a new tool based on peer-to-peer IF-TOPSIS, intended for micro-level assessment of medical devices with the main objective of overcoming the above-mentioned critical issues and provide local institutions/hospitals with a general tool. An example of application of this tool in the choice between three neurological devices is shown. Contributions are both theoretical and practical. Theoretically, while proposing a general MCDA tool for micro-level HTA, we answer the call for the identification of key methodological principles for the local assessment of medical devices. At the practical level, this tool is readily implementable and can be adapted to consider the local idiosyncratic characteristics of the context where decisions have to be made.

KEYWORDS

health technology evaluation, IF-TOPSIS, MCDA, peer-group decision making

1 | INTRODUCTION

The progressive improvements of scientific research in the development of technologies embedded in medical devices is reflected in the continuous growth of the rates of introduction of these technologies in the health sector (Chandra & Skinner, 2012). These innovative medical devices, while effectively responding to the population's growing needs in terms of disease diagnosis and treatment, need to be evaluated through a decisional process capable to take into account a broad range of criteria, such as clinical value, safety, potential and operational effectiveness, economic and organisational impact (Tarricone, Callea, et al., 2017a), as well as the burdens the end user suffers (Martelli et al., 2016).

Health Technology Assessment (HTA) allows to effectively tackle such complex decision process. Indeed, HTA, through the use of Multi-Criteria Decision Analysis (MCDA), can include a multiplicity of assessment aspects that are not necessarily linked to each other or that even conflict the one with the other (Adunlin et al., 2014; Lazzini, 2014). Hence, in recent years, experiments on approaches that combine HTA and MCDA (Martelli et al., 2016) have been widely investigated, in both public and private areas. Considerable research has been done in the field of micro-assessment (Sloane, 2004), also referred to as micro-level HTA (Favaretti et al., 2009), or local/hospital-based HTA (Gagnon et al., 2014), to support local decision makers in selecting potentially very different medical devices (Sampietro-Colom et al., 2012).

However, the literature documents that locally led HTA tools which build on MCDA are very different the one from the others, especially in terms of the utilized criteria that indeed are strictly tailored to the specific technology to be evaluated (Sloane, 2004). Albeit partially explainable with the necessity to cater to specific needs of the contexts in which these tools are to be implemented (Ehlers et al., 2006; Gagnon et al., 2014), such overwhelming diversity between the local HTA tools is due to a lack of clear guidance at the (inter)national level. This is because, firstly, only a small proportion of devices used in hospitals are actually assessed by national health authorities (Martelli et al., 2017). Secondly, even when conducted at national or international levels, HTA of medical devices has been largely developed with the evaluation of drugs/pharmaceuticals in mind (Ciani et al., 2017), so overlooking important differences between drugs/pharmaceuticals and devices, and neglecting pivotal aspects of medical devices.

This way, the practical implementation of HTA tools has been seriously jeopardised, since decision makers are required to identify new criteria each time a new device has to be assessed, propose adequate measurement scales for them, and finally validate them.

This paper, while answering the call for a general tool for local HTA (Martelli et al., 2017), proposes a set of criteria which apply to any device (Sampietro-Colom et al., 2012). The suitability of the tool is shown through its application in an hospital that had to choose between three devices addressed to people that, because affected by brain damage of ischemic/traumatic origin, had communication problems. A sensitivity analysis is also provided to ensure the developed evaluation tool's internal validity.

This article is structured as follows: next section presents the analysis of the extant scientific literature on hospital-based HTA of medical devices; Section 3 describes the tool we propose with its criteria; Section 4 shows the methodology and its application in the analyzed case study (analysis of the decision making problem, the alternatives and the result of MCDA application). Finally discussion and conclusion are given.

2 | LITERATURE ANALYSIS

The number of hospitals performing local HTA has grown around the world (Favaretti et al., 2009; Gagnon, 2014; Sampietro-Colom et al., 2012) and this nurtured the creation of a plethora of different local HTA tools. If, on the one hand, this variety of tools is due to the necessity of taking into account the idiosyncratic characteristics of the local context (Sampietro-Colom & Martin, 2016) – for example, the innovativeness of the technology in teaching hospitals (Sampietro-Colom et al., 2012) - on the other hand, this is due to the absence of a high-level common methodological HTA guidance (Martelli et al., 2017). Besides, this inconsistent development has been exacerbated by the hospitals' need to conduct micro-level HTA on those devices which are disregarded at the (inter)national level, because not implantable (Martelli et al., 2017), not big/tickets (Sampietro-Colom et al., 2012), or characterised by a low level of risk (Fuchs et al., 2017): while having a considerable weight on hospitals' corporate financial statements, these tools need to be assessed, also in the light of the growing concerns about limited resources (Barron et al., 2015; Ivlev et al., 2015).

Therefore, urged by the need to evaluate medical devices and without any guidance, local institutions/hospitals had to decide between standard tools – for example, Incremental Cost-Effectiveness Ratio (ICER) or QALY (Incremental Cost for Quality-Adjusted Life-Year - Thokala & Duenas, 2012; Oliveira et al., 2019; Marsh et al., 2014) - or tailored tools able to take into account the multidimensional and multidisciplinary aspects that pertain to the healthcare decision-making processes. Albeit implicitly, this choice entails a trade-off in the number of criteria: standard tools build on only two criteria (Devlin & Sussex, 2011; Marsh et al., 2014; Wahlster et al., 2015), while tailored tools include a wide spectrum of criteria to be conceived whenever a new device is to be assessed.

In many cases, local institutions opted for tailored tools able to include a large set of criteria, and adopted MCDA techniques to get to the final decision. Indeed, firstly, this choice allows local institutions to overcome the pitfalls ascribable to single indicators (Wahlster et al., 2015), such as the exclusion of socially-valuable benefits that go beyond the improvements in length and health-related quality of life (Devlin & Sussex, 2011; Marsh et al., 2014). Secondly, this choice allows local institutions to cope with the multiplicity of factors which impact health decisions, the high number of alternatives and the possible trade-offs between them, the imperfect information, and the high levels of complexity typical of HTA decisions (Mathew, 2011; Thokala et al., 2016). Thirdly, MCDA-based HTA prevents the decision-making process from being based on principles, values and processes that are not explicitly identified (Devlin & Sussex, 2011): this way, decisions take the distances from the historical approach that, while being implicit on the relevant criteria, led to a non-transparent and non-replicable decision-making (Sampietro-Colom et al., 2012).

However, local tools developed inconsistently, building on criteria specifically bespoken for the specific technology to be evaluated (Ottardi et al., 2017; Sloane, 2004; Vettoretto et al., 2018; Wahlster et al., 2015). This way, criteria identification followed a bottom-up approach (Marsh et al., 2016), which entails that criteria are '*very specifically relevant to the problem at hand, but* […] *are not 'portable' to new decision situations*' (Morton & Fasolo, 2009; p. 269). What is more, differences are not only between the set of criteria, but also within the same criterion. For instance, the economic impact ranges from costs to cost-effectiveness, from impact on other spending to incremental cost-effectiveness and affordability (Wahlster et al., 2015). This approach, while affecting generalizability and structuring of the criteria, negatively impacts on the timeliness and consistency of the decisions, respectively. In this regard, the impossible application of criteria to devices other than those for whom criteria were conceived (lack of generalizability) implies identifying new criteria, anytime a device ever assessed before has to be evaluated. This requires that, before the very assessment can be implemented (Blythe et al., 2019; Bretoni et al., 2019), criteria are designed and validated to check for their plausibility (Marsh et al., 2016). In its turn, this lack of criteria generalizability affects decisions' timeliness, fundamental in healthcare contexts where the timely use of a new technology can be decisive for diagnosing or accessing treatment. In addition, the lack of the same set of consolidated criteria between different evaluations has immediate implications on the consistency of the decisions. Indeed, using different sets of criteria each time a device has to be assessed may lead to problems in the structuring phase of the measurement model: the unintended double counting of one/more criteria (Baltussen et al., 2019) (typical of criteria connected with cost, Wahlster et al., 2015) as a result of lack of mutual exclusivity (Devlin & Sussex, 2011), the overlapping between criteria (Pecchia et al., 2019) or, conversely, the unintentional exclusion of one/more criteria in the assessment of a device (Baltussen et al., 2019). The consequences in terms of inconsistency from decision to decision are immediate. Last but not least, claims of subjective interpretations of newly designed criteria may be raised (Oliveira et al., 2019). As a final result, the decision makers' credibility and potentially legitimacy may be questioned (Thokala et al., 2016).

Given these premises, a few articles have begun to highlight the need to create a decision support system for the selection of medical devices (Marsh et al., 2014) that takes into account a wide and stable set of criteria (Wahlster et al., 2015), based on clear and above all shared assessments (Ivlev et al., 2015). Our article takes up this challenge and builds on criteria that, despite not gathered in a single tool, are provided by both the scientific literature and Regulative Agencies. Specifically, these criteria can be gathered in three main macro-areas: medical-sanitary, techno-engineering, and sustainability. Batavia and Hammer (1990) and Abdel-Basset et al. (2019), while attributing importance to the medical-sanitary aspects, agree with Mathew (2011) in considering the device's levels of efficacy and efficiency, but also add the degree of satisfaction with the use of the device, the patient's comfort, the versatility of the device to the type of patient and its interoperability. Furthermore, Batavia and Hammer (1990) add the compatibility of the sterilisation method, while Martelli et al. (2016) the clinical impact, in terms of a possible reduction in the patient's disability following the use of the medical device.

Mathew (2011), Batavia and Hammer (1990), and Martelli et al. (2016) also agree on the need to consider the medical-sanitary area, with criteria on the level of safety and innovativeness of the device. Regarding this latter aspect, Regulative Agencies such as AIFA (the Italian Drugs Agency) and ASSOBIOMEDICA (Italian General Confederation of the Medical Device Industry) add further criteria in their guidelines. For example, AIFA suggests a criterion for evaluating the satisfaction of a therapeutic need and the added therapeutic/technological value, while ASSOBIOMEDICA indicates considering the period of technological adequacy.

The literature also addresses device sustainability, proposing a set of criteria for its evaluation - be it in economic-managerial, social, or environmental terms. The economic sustainability is the mostly addressed issue in the literature: Park et al. (2019) and Abdel-Basset et al. (2019) agree in considering the cost of the device, Mathew (2011) suggests using the output of the cost-effectiveness analysis, while Martelli et al. (2016) adds the Diagnostic Related Group. Costs related to device maintenance, operators' training and acquisition of the skills necessary for device management are also found in many contributions. Greater and greater attention is devoted to possible indirect costs patients may bear (Batavia and Hammer (1990)).

The technical-engineering criteria used so far are extremely heterogeneous. For example, we can find the criteria specifically tailored for the evaluation of the Magnetic Resonance Imaging (Ivlev et al., 2015), Chronic Obstructive Pulmonary Disease (Marsh et al., 2017), 3d and 2d laparoscopy (Vettoretto et al., 2018), reusable pedicle screw tool kits for lumbar arthrodesis (Ottardi et al., 2017), and neonatal ventilators (Sloane, 2004). Anyhow, there are also more general technical-engineering criteria; for instance, Batavia and Hammer (1990) proposes the adequacy of the dimensions and weight of the device, the compatibility of the material, the time of use, and the alarm system.

This paper, while building on the suggestions of the literature, aims at identifying a viable tool for the local HTA of medical devices that avoids bespoke criteria and proposes a set of criteria which is usable for all medical devices, whatever their class. Hence this paper can be considered a first step toward the definition of a general procedure, that at the same time allows elements of flexibility to take into account the idiosyncratic characteristics of the context (Sampietro-Colom et al., 2012; Sampietro-Colom & Martin, 2016). Given the prominent withdrawals and problems of devices over the past decade, we deem this tool is very timely and useful.

3 | THE HTA TOOL AND ITS CRITERIA

For defining criteria suitable for any medical device, whatever their class, we analyzed two main sources: on the one hand, multiple datasheets of medical devices of each class, and, on the other, the scientific literature.

As regards datasheets, the following were analyzed: AAET for Italian Ministry of Health (2016), Medical devices datasheets from FDA, U.S. Food, & Drug Administration, Medical Device Databases, n.d., European Commission Enterprise and Industry Directorate General, Guidelines on Medical Device (2009). As regards the analysis of the extant literature, Scopus and PubMed databases were used, in a time span ranging from 2000 to 2018, with the following keywords: (HTA for Medical Devices Rating) OR (MCDA for HTA) OR (MCDA for Technology Assessment) OR (MCDA for Medical Devices) OR (Criteria for HTA) OR (Criteria for Technology Rating) OR (Criteria for Medical Devices Rating) OR (Evaluation of Medical Devices) OR (Costs Criteria for Medical Devices Evaluation).

The criteria were validated by specialists from the HTA area of the Regional Technical and Administrative Support Agency, and from the Pharmaceutical and Medical Devices Management Unit of the University Hospital. According to the literature, three macro-areas were identified. Every macro-area is divided into areas, each containing a number of criteria. Specifically:

(a) Techno-engineering macro-area: it concerns the technical criteria mainly emerged from the analysis of the datasheets. This macro-area includes the following areas: appearance, composition, durability, technical features, measurement, power supply, and alarm systems;

(b) Medical-Sanitary macro-area: it includes the clinical criteria aimed at ensuring that the supported therapeutic/diagnostic procedure achieves a good performance. It included the following areas: treatment, use, field of application, clinical impact, and innovativeness;

(c) Sustainability macro-area: it regards the domain of economic/managerial, social and environmental sustainability.

Table 1 provides greater details on the areas belonging to each macro-area, all the identified criteria and the connected references. Not all the criteria are applicable to medical devices of any class; for this reason, it was decided to consider two degrees of freedom: one concerning the possibility of deleting criteria which do not apply to the device being studied or to the specific decision-making context, the other concerning the possibility of adding new criteria, specific to the medical device being studied or the decision-making context.

4 | METHODOLOGY AND APPLICATION

4.1 | The decision-making problem

Our study regards the selection of an assistive medical device within an organisation of the Italian National Health Service (NHS) - at the top of the world rankings for efficiency, life expectancy, as well as relative and absolute costs (Miller & Wei, 2018) - with its high number of public and private investments observed by the National Anti-Corruption Authority (ANAC). The Italian legislation classifies medical devices according to four classes (I, IIa, IIb, III), according to the type of power supply, invasiveness, duration of use, potential, and risk associated with its use.

Specifically, the decision-making context concerns the selection of a medical device addressed to people with communication problems, being affected by brain damage of ischemic/traumatic origin. Communication problems are indeed particularly felt by locked-in state (LIS) or completely LIS (CLIS) patients, no longer able to use eye-tracking for communicating because of the ocular muscles paralysis. CLIS patients hardly benefit from the assistive technologies today on the market.

Three devices supporting both LIS and CLIS patients were identified. The first (A) is a class I device based on brain-computer Interface (BCI) technology capable of converting the brain electrical signals into outputs that control external devices, without using voluntary movements. The second (B) is a IIb class device, based on BCI technology, that uses electroencephalography sensors; it is non-invasive and builds on the Visually Evoked Potential. The third (C) belongs to class III; it requires the implantation of a chip between the brain and the skull. As can be seen, the three devices differ considerably from each other in terms of power supply, invasiveness, duration of use, potential, and the risk associated with use. Indeed, as often happens in device selection, very different devices belonging to different classes cater to the same need. Thus, in a situation of ambiguity, uncertainty, and vagueness, the decision-makers are called upon to evaluate three medical devices with very different clinical, technical, and sustainability characteristics but used for the same purpose. As an addition, decision-makers, unlike what happens for drugs, do not have clinical studies and trials on medical devices (Akpinar et al., 2019) and this exacerbates ambiguity and uncertainty (Mathew, 2011).

In the investigated case, the devices' assessment has been assigned to five professionals: three are highly knowledgeable professionals in the neurological area, with proven and long clinical practise and management experience of complex structures in hospitals, and two senior administrative managers, experienced in the governance of the NHS' Healthcare Companies. Specifically, they are:

- DM1: Physician of the Neurology/Stroke Unit;
- DM2: Director of the Neurology/Stroke Unit;
- DM3: Emeritus Director of the Neurology/Stroke Hospital Unit; Past President of the Scientific Society of Neurologists, Neuroradiologists and Hospital Neurosurgeons (S.N.O.);
- DM4: Administrative Director of the NHS Organisation at which the Neurology/Stroke Unit belongs;
- DM5: General Manager of the NHS Healthcare Organisation at which the Neurology/Stroke Unit belongs.

Between the decision-makers, each acquainted with the abilities and experience of the others, there are not subordination relationships. This requires that in the implementation of the MCDA tool - preliminarily to the evaluation of the relevance (weights) of the criteria and of the impact of the different devices on the provided criteria - a peer procedure for determining the weights of the decision-makers' opinions is used.

TABLE 1 Criteria and selected criteria (*)

Macro-area	Area	Criteria	Description	Sources
Techno-engineering properties	Appearance	Dimensions adequacy* Weight adequacy*	Dimension and weight can influence the use of device (e.g., its handling, bulkiness, etc.).	Batavia and Hammer (1990) Medical devices datasheets from AAET for Italian Ministry of Health, Medical Device Data Bank (2016) Medical devices datasheets from FDA, U.S. Food, and Drug Administration, Medical Device Databases (n.d.) European Commission Enterprise and Industry Directorate General, Guidelines on Medical Device (2009)
	Composition	Material compatibility	The material of the device can be a concern in case of allergies or intolerances, when the device is used by physicians on the patient.	
		Surface treatment compatibility	The surface treatment of devices metal components can influence the interaction between the different components of device and/or with other metal devices.	
	Durability	Life-time adequacy*	The life-time adequacy assesses if the limit within which it'is possible to use the device (before it loses or changes its physical, chemical or pharmaceutical properties) is adequate respect to the clinical case.	
	Technical features	Strain Malleability Strength Transpiration …	One or more of these technical features can be decisive for specific uses of medical devices in specific medical processes.	
	Measurement	Number of detected parameters*	The number of detected parameters indicates how many measures the tool can record.	
		Adequacy of the measurement method approximation	The type of measurement and its approximation can influence the correct display of the clinical values measured by the device.	
		Adequacy of the measurement range	The range of measurement assesses if the minimum and maximum values measured by the device are adequate respect to the clinical case (e.g., in devices designed to measure the weight of severely obese patients).	
	Power supply	Power supply type compatibility*	The compatibility of power supply type (battery, pulsed power supply, continuous feeding) of the device can influence the diagnosis process, the treatment of the pathology and the possibility to move the device.	
		Battery recharge time compatibility*	The time for battery recharge can influence the performance of the treatment/diagnosis procedure due to periods to be waited before the device can be reused.	
		Electric consumption* Electric classification level Absorbed power	Electric consumption, electric classification level and absorbed power affect saving of energy.	
	Alarm system	Alarm accuracy*	This criterion evaluates the need of a greater alarm accuracy for special device rooms (e.g., Computer Axial Tomography, CAT, room).	
		Number of alarms*	The number of alarms indicates how many types of alarm the device has (e.g., sonorous, visible…).	
Medical-Sanitary Properties	Treatment	Compatibility of sterilisation treatment	The type of sterilisation used by the producer in the production phase can influence the re-sterilisation process and hence the possibility of re-use of the device.	Abdel-Basset et al. (2019) AIFA, Agenzia Italiana del Farmaco, Determina n. 1535/2017, Allegato 1 'Criteri per la valutazione dell'innovatività' (2017)
	Use	Effectiveness level*	The effectiveness level quantifies the extent to which the device caters the user's clinical need.	

(Continues)

TABLE 1 (Continued)

Macro-area	Area	Criteria	Description	Sources
		Efficiency level*	The efficiency level quantifies the length of the time taken to achieve the target.	ASSOBIOMEDICA - Studi numero 38 - novembre 2017 osservatorio parco installato: le apparecchiature di diagnostica per immagini in Italia (2017) Batavia and Hammer (1990) Martelli et al. (2016) Mathew (2011)
		Satisfaction*	This criterion evaluates the satisfaction of the physician in the use of the device.	
		Patient comfort	This criterion evaluates how comfortable is for the patient the use of the device.	
		Duration of each application	The duration of each device application can influence the scheduling of hospital activities.	
		Interoperability degree	This criterion indicates the new device's capabilities to satisfy several operating units' needs and integrate with technologies already in use.	
	Field of application	Versatility with respect to the types of patients*	It indicates the possibility to use the device for different types of patients (e.g., differentiated by age: adults, children; males, females; ...).	
		Versatility with respect to the room temperature	It indicates the possibility to use the device in different laboratories and clinical spaces where the room temperature can influence its use (e.g., in the cryogenics laboratories).	
	Clinical impact	Reduction in drug consumption*	The use of the device can allow a reduction of drug use (above all in terms of pharmaceutical cost).	
		Disability reduction*	The use of a device can significantly reduce the disability of patients.	
	Innovativeness	Fulfilment of a therapeutic need*	This criterion evaluates whether the technology is capable to give response to therapeutic needs of population not yet resolved by other devices. This is fundamental innovation, capable of satisfying a need to which before there was no response or capable of catering a need in a new way.	
		Added therapeutic/technological value*	Added therapeutic/technological value considers the entity of the clinical benefit offered by the new technology respect to the available alternatives (if existing). This is functional innovation, that is, the addition of functionalities which are not present in competitors' devices.	
		Technological adequacy period	This criterion indicates the time-in-years limit beyond which the medical device is obsolete.	
		Software break risk*	This criterion evaluates the probability of the device software breaking occurrence and its impact on the result.	
		Hardware break risk*	This criterion evaluates the probability of the device hardware breaking occurrence and its impact on the result.	
		Risk of infection	The criterion evaluates the probability of posthumous infection due to the use of device and its impact on user health and on compensation costs.	
		Degree of security	The security degree evaluates how safe is the device use, that is, how the device follows the standards for the safety of medical devices (EN ISO 14971).	
Sustainability	Economic and managerial	DRG (revenue)	DRG (diagnosis-related group) evaluates the company revenue for each diagnostic/surgical procedure within which the device is used.	Abdel-Basset et al. (2019) Batavia and Hammer (1990) Martelli et al. (2016) Mathew (2011) Park et al. (2019)
		Device cost*	It is the cost of the device.	
		Cost of the diagnostic/surgical procedure*	This criterion evaluates the cost of the diagnostic/surgical procedure within which the device is used (e.g., cost of personnel used, cost of other auxiliary equipment...)	

TABLE 1 (Continued)

Macro-area	Area	Criteria	Description	Sources
		Trial cost*	This criterion evaluates the cost of the trial before the device installation.	
		Post-diagnostic/surgical procedure costs	This criterion evaluates the cost of the post-diagnostic/surgical procedure after the medical process within which the device in used.	
		Compatibility with the purpose of the structure	It is the compatibility between the device and the purpose of the structure (clinical unit) which is addressed.	
		Job reconfiguration	The use of a new device could need the reconfigurations of the staff tasks.	
		Cost of personnel skills adaptation	This criterion evaluates the cost of adaptation of personnel skills as a consequence of the device introduction.	
		Cost of infrastructure adaptation	This criterion evaluates the cost of infrastructure adaptation following the device introduction.	
		Cost of engineering support*	This criterion evaluates the cost of engineering support following the device introduction.	
		Reuse costs	This criterion evaluates the cost to reset the device for its reuse.	
		Cost of drugs	This criterion evaluates the cost of pharmaceuticals use in pre and/or post clinical process within which the device is used.	
		Depreciation rate unitary cost	This criterion evaluates the medical device's economic flow, calculated as the annual depreciation rate on yearly performances.	
		Leasing rate unitary cost	This criterion indicates how much the medical device's leasing weighs in the long-term.	
		Maintenance unitary cost	This criterion indicates how much the maintenance cost is distributed for each service performed with the medical devices.	
		Quality cost	It is the sum of the costs to prevent the occurrence of non-conformities of the diagnosis or treatment according to ISO 8402 standard.	
		Weight on total investments	It is an index of how much the single investment weighs on the entire company's core business and calculates as the Return On Investment (ROI) variation after the new investment.	
	Social	Patient's indirect costs*	The patient's indirect costs consider all costs which burden the end user (e.g., cost of health homecare, transport cost to and from the hospital…).	
	Environmental	Battery sustainability*	The battery sustainability regards the management and disposal issues of the device battery.	

As regards the MCDA technique, our choice fell on peer-to-peer IF-TOPSIS (Aloini et al., 2014, 2017) which meets the following needs emerged during the context analysis: firstly, it is compliant with the Italian anti-corruption legislation. More in particular, ANAC's code of conduct provides three possible multi-criteria techniques to choose from for medical devices purchasing: AHP, ELECTRE, and TOPSIS. Secondly, it is able to cope with the ambiguity, uncertainty, and vagueness of the specific decision-making contexts. Thirdly, it takes into account the composition of the commission charged with the responsibility of assessing medical devices; the commission indeed is composed of peers at the same organisational level, without no possibility to identify a super-decision maker.

The criteria for the implementation of the MCDA method were selected from the tool proposed in Section 3; they are marked by * in Table 1.

The five decision-makers were asked to evaluate (1) the importance of each criterion, (2) the relevance of the judgement of each decision-maker, (3) the impact of different technological alternatives on criteria. For (1) and (2), linguistic variables measured on a five-level scale (from Very Important to Very Unimportant) were adopted. For (3), a Likert Scale from 1 to 4 was used (except for the criteria 'Length of lifetime', whose scale ranges from 1 to 3).

4.2 | MCDA implementation

This technique consists of eight steps and leads to a ranking of the possible choices, ordered according to a decreasing proximity coefficient (C_i), which measures how much each alternative (i) differs from the ideal solution. The steps, whose implementation is detailed in the Data S1 (A.1.), are:

I. Construction of the aggregated importance IF decision matrix: after each decision-maker has voted the importance of the others' opinions and their own, this step's main output is merging all decision-makers' opinions into an opinion group.

II. Calculation of the weight of each decision-maker: the individual opinions are merged into a whole judgement to determine a grade of decision-makers' importance range (because all decision-makers have not the same prominence).

III. Construction of the aggregated IF decision matrix: each alternative is evaluated by the decision-makers using the selected criteria, and the aggregation of all the opinions builds a decision matrix.

IV. Calculation of the weights of the single criterium: the individual opinions are fused into a whole judgement to determine a grade of criteria importance range (because all criteria have not the same prominence).

V. Construction of the aggregated weighted IF decision matrix: while building on III and IV steps' outputs, the aggregated weighted intuitionistic fuzzy decision matrix is constructed.

VI. Calculation of the intuitionistic fuzzy positive-ideal solution and the intuitionistic fuzzy negative-ideal solution.

This step involves dividing the criteria between benefits and costs. This step allows determining for each alternative the intuitionistic fuzzy positive-ideal solution and the negative-ideal solution one.

VII. Calculation of the separation measures and the proximity coefficients. With each alternative's separation measures, calculated using the normalised Euclidean distance from the VI step outputs, the relative proximity coefficients to the ideal solution are achieved.

VIII. Rank of alternatives according to the decreasing proximity coefficient C_i. In this last step, by decreasingly ordering the relative proximity coefficients, the possible alternatives are ranked.

Following these steps, the proximity coefficients obtained for each of the three alternatives are $C_A = 0.687$, $C_B = 0.393$, and $C_C = 0.301$.

Therefore, the ranking of the alternatives is as follows: A – B – C.

To ensure the developed evaluation tool's internal validity, the following sensitivity analysis of the results was performed. This allowed coping with the biases and limitations of the method and determine the robustness of HTA findings and conclusions (Drummond et al., 2008).

We proceeded to examine the dispersion of each possible alternative's proximity coefficient (C_{in}) – the dependent variable – as the decision-makers 'Opinion Value' coefficient (ϕ_l) – the independent variable – varies.

We have chosen ϕ_l as an independent variable because its value (which expresses the calibre of each decision-makers' opinion) contributes to determining D_k, that is, the aggregate importance of the k-th decision-maker, used to compute the weights of the evaluation criteria and the evaluators' judgement. These ones influence the values of C_i through which the ranking to support the choice is made. Therefore, the unweighted choice of ϕ_l could distort the entire evaluation process.

The ϕ_l value was varied in the range [0.1:0.1:0.6] for each of the five decision-makers, considering 0.1 the minimum value of ϕ_l, and 0.6 the maximum one (with unit of step equal to 0.1), and the sum of the individual opinions' values must be equal to 1.

One hundred twenty-six different possible combinations of the values of the DMs' opinion that influence the evaluation criteria importance on the final solution are obtained. Table 2 shows the n_{th}-combinations of the DMs' 'Opinion Value' coefficients and the relative i_{th}-proximity coefficients obtained. Figure 1 shows the proximity coefficient's dispersion obtained in the 126 different combinations of the DMs opinion values for each of the three alternatives.

The proximity coefficients as the opinion value changes are ordered, starting from the left, by the configuration with the utmost importance of the opinion of the decision makers of the administrative-governance area ($\phi_1 = 0.1$, $\phi_2 = 0.1$, $\phi_3 = 0.1$, $\phi_4 = 0.1$, $\phi_5 = 0.6$) to the configuration, increasing to the right, with the utmost importance of the opinion of the clinical decision makers ($\phi_1 = 0.6$, $\phi_2 = 0.1$, $\phi_3 = 0.1$, $\phi_4 = 0.1$, $\phi_5 = 0.1$).

The red points indicate the proximity coefficients values obtained considering the configuration in which the DMs have the same importance; the opinion values are equal ($\phi_1 = 0.2$, $\phi_2 = 0.2$, $\phi_3 = 0.2$, $\phi_4 = 0.2$, $\phi_5 = 0.2$), which is the hypothesised combination choice in the presented application of the evaluation method.

The results of the descriptive analysis of the variable C_{in} are shown in Table 3. From this, multiplying by 100 the standard deviation divided by the mean - both computed, for each C_{in} - the following values of the coefficients of variation of each alternative's proximity coefficient (CV_i) are obtained:

$$CV_A = 0.155,$$

$$CV_B = 0.425,$$

$$CV_C = 0.758$$

The CV of A, the first medical device in the alternatives' ranking, is the lowest. Besides, the ranking CV_A – CV_B – CV_C follows the hierarchy of the alternatives identified with the implementation of the method developed before (A – B – C). This is a satisfactory proof of the method's robustness in determining the best solution and the composition of the alternatives ranking.

TABLE 2 Variation of the proximity coefficients C_{in} as the ‘Opinion Value’ coefficients vary

	‘Opinion Value’ coefficients ϕ_i					Proximity coefficient C_{in}		
n-combinations	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
1	0.1	0.1	0.1	0.1	0.6	0.687101	0.392765	0.303597
2	0.1	0.1	0.1	0.2	0.5	0.68598	0.394335	0.305533
3	0.1	0.1	0.1	0.3	0.4	0.686978	0.392555	0.304805
4	0.1	0.1	0.1	0.4	0.3	0.685322	0.395374	0.306596
5	0.1	0.1	0.1	0.5	0.2	0.685325	0.39562	0.306526
6	0.1	0.1	0.1	0.6	0.1	0.685749	0.396599	0.305537
7	0.1	0.1	0.2	0.1	0.5	0.687831	0.391923	0.302272
8	0.1	0.1	0.2	0.2	0.4	0.686777	0.393504	0.304108
9	0.1	0.1	0.2	0.3	0.3	0.686343	0.394238	0.304828
10	0.1	0.1	0.2	0.4	0.2	0.686251	0.394669	0.304921
11	0.1	0.1	0.2	0.5	0.1	0.68668	0.395771	0.303969
12	0.1	0.1	0.3	0.1	0.4	0.687981	0.391783	0.302004
13	0.1	0.1	0.3	0.2	0.3	0.686957	0.393377	0.303783
14	0.1	0.1	0.3	0.3	0.2	0.686607	0.394215	0.304321
15	0.1	0.1	0.3	0.4	0.1	0.68694	0.395489	0.303537
16	0.1	0.1	0.4	0.1	0.3	0.68804	0.391757	0.30189
17	0.1	0.1	0.4	0.2	0.2	0.687104	0.393445	0.303487
18	0.1	0.1	0.4	0.3	0.1	0.687185	0.395104	0.303129
19	0.1	0.1	0.5	0.1	0.2	0.688184	0.391777	0.301595
20	0.1	0.1	0.5	0.2	0.1	0.687702	0.394231	0.302256
21	0.1	0.1	0.6	0.1	0.1	0.688897	0.39224	0.300148
22	0.1	0.2	0.1	0.1	0.5	0.68754	0.393699	0.302411
23	0.1	0.2	0.1	0.2	0.4	0.68648	0.395236	0.304303
24	0.1	0.2	0.1	0.3	0.3	0.686046	0.395947	0.305033
25	0.1	0.2	0.1	0.4	0.2	0.685961	0.396384	0.305127
26	0.1	0.2	0.1	0.5	0.1	0.686415	0.397566	0.304129
27	0.1	0.2	0.2	0.1	0.4	0.688161	0.392909	0.30126
28	0.1	0.2	0.2	0.2	0.3	0.687185	0.394466	0.303015
29	0.1	0.2	0.2	0.3	0.2	0.686865	0.395302	0.303531
30	0.1	0.2	0.2	0.4	0.1	0.687247	0.396679	0.302693
31	0.1	0.2	0.3	0.1	0.3	0.688313	0.392787	0.300978
32	0.1	0.2	0.3	0.2	0.2	0.68744	0.394461	0.302528
33	0.1	0.2	0.3	0.3	0.1	0.687588	0.396217	0.302108
34	0.1	0.2	0.4	0.1	0.2	0.688481	0.392939	0.300644
35	0.1	0.2	0.4	0.2	0.1	0.688094	0.395391	0.301224
36	0.1	0.2	0.5	0.1	0.1	0.689219	0.393536	0.299176
37	0.1	0.3	0.1	0.1	0.4	0.687765	0.394093	0.301893
38	0.1	0.3	0.1	0.2	0.3	0.686739	0.395632	0.303732
39	0.1	0.3	0.1	0.3	0.2	0.686402	0.396401	0.304274
40	0.1	0.3	0.1	0.4	0.1	0.68678	0.397841	0.303425
41	0.1	0.3	0.2	0.1	0.3	0.688362	0.393338	0.300768
42	0.1	0.3	0.2	0.2	0.2	0.687497	0.395001	0.302319
43	0.1	0.3	0.2	0.3	0.1	0.687662	0.396784	0.301882
44	0.1	0.3	0.3	0.1	0.2	0.688619	0.393302	0.300272
45	0.1	0.3	0.3	0.2	0.1	0.688261	0.395881	0.300828

(Continues)

TABLE 2 (Continued)

n-combinations	'Opinion Value' coefficients ϕ_i					Proximity coefficient C_{in}		
	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
46	0.1	0.3	0.4	0.1	0.1	0.689384	0.394065	0.298767
47	0.1	0.4	0.1	0.1	0.3	0.68789	0.394266	0.301623
48	0.1	0.4	0.1	0.2	0.2	0.686962	0.395919	0.303266
49	0.1	0.4	0.1	0.3	0.1	0.687102	0.39768	0.302845
50	0.1	0.4	0.2	0.1	0.2	0.688582	0.393606	0.300299
51	0.1	0.4	0.2	0.2	0.1	0.688223	0.396189	0.300859
52	0.1	0.4	0.3	0.1	0.1	0.689436	0.394296	0.298635
53	0.1	0.5	0.1	0.1	0.2	0.688078	0.39441	0.301237
54	0.1	0.5	0.1	0.2	0.1	0.687652	0.396956	0.301873
55	0.1	0.5	0.2	0.1	0.1	0.689366	0.394486	0.298747
56	0.1	0.6	0.1	0.1	0.1	0.688887	0.395142	0.299661
57	0.2	0.1	0.1	0.1	0.5	0.687681	0.391809	0.30184
58	0.2	0.1	0.1	0.2	0.4	0.686645	0.393399	0.303784
59	0.2	0.1	0.1	0.3	0.3	0.686214	0.394131	0.304536
60	0.2	0.1	0.1	0.4	0.2	0.686118	0.394556	0.30464
61	0.2	0.1	0.1	0.5	0.1	0.686505	0.39562	0.303656
62	0.2	0.1	0.2	0.1	0.4	0.68845	0.391081	0.300643
63	0.2	0.1	0.2	0.2	0.3	0.68748	0.392689	0.302454
64	0.2	0.1	0.2	0.3	0.2	0.68715	0.393528	0.302993
65	0.2	0.1	0.2	0.4	0.1	0.687479	0.394799	0.302168
66	0.2	0.1	0.3	0.1	0.3	0.688631	0.390968	0.300354
67	0.2	0.1	0.3	0.2	0.2	0.687755	0.392673	0.301959
68	0.2	0.1	0.3	0.3	0.1	0.687851	0.394339	0.301564
69	0.2	0.1	0.4	0.1	0.2	0.688794	0.390996	0.300017
70	0.2	0.1	0.4	0.2	0.1	0.68836	0.393484	0.300654
71	0.2	0.1	0.5	0.1	0.1	0.689463	0.391565	0.298547
72	0.2	0.2	0.1	0.1	0.4	0.688095	0.392935	0.30075
73	0.2	0.2	0.1	0.2	0.3	0.687122	0.394481	0.302615
74	0.2	0.2	0.1	0.3	0.2	0.686797	0.395306	0.303163
75	0.2	0.2	0.1	0.4	0.1	0.687144	0.396645	0.302301
76	0.2	0.2	0.2	0.1	0.3	0.688781	0.39225	0.299672
77	0.2	0.2	0.2	0.2	0.2	0.687948	0.393915	0.301247
78	0.2	0.2	0.2	0.3	0.1	0.688093	0.395655	0.300809
79	0.2	0.2	0.3	0.1	0.2	0.689049	0.392228	0.299186
80	0.2	0.2	0.3	0.2	0.1	0.68869	0.394774	0.299758
81	0.2	0.2	0.4	0.1	0.1	0.689758	0.392985	0.297683
82	0.2	0.3	0.1	0.1	0.3	0.688324	0.393407	0.300231
83	0.2	0.3	0.1	0.2	0.2	0.687453	0.395051	0.301884
84	0.2	0.3	0.1	0.3	0.1	0.687586	0.396791	0.301446
85	0.2	0.3	0.2	0.1	0.2	0.689074	0.392832	0.29899
86	0.2	0.3	0.2	0.2	0.1	0.68873	0.395389	0.299554
87	0.2	0.3	0.3	0.1	0.1	0.689888	0.393568	0.297346
88	0.2	0.4	0.1	0.1	0.2	0.688549	0.39369	0.299741
89	0.2	0.4	0.1	0.2	0.1	0.688155	0.396223	0.300367
90	0.2	0.4	0.2	0.1	0.1	0.689839	0.393889	0.297355

TABLE 2 (Continued)

	'Opinion Value' coefficients ϕ_i					Proximity coefficient C_{in}		
n-combinations	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
91	0.2	0.5	0.1	0.1	0.1	0.689294	0.394554	0.298123
92	0.3	0.1	0.1	0.1	0.4	0.687943	0.391419	0.300921
93	0.3	0.1	0.1	0.2	0.3	0.686983	0.392999	0.302813
94	0.3	0.1	0.1	0.3	0.2	0.686645	0.393824	0.30338
95	0.3	0.1	0.1	0.4	0.1	0.686922	0.395066	0.302559
96	0.3	0.1	0.2	0.1	0.3	0.688677	0.390744	0.299741
97	0.3	0.1	0.2	0.2	0.2	0.687844	0.392433	0.301354
98	0.3	0.1	0.2	0.3	0.1	0.68793	0.394093	0.300952
99	0.3	0.1	0.3	0.1	0.2	0.688942	0.390699	0.299239
100	0.3	0.1	0.3	0.2	0.1	0.688535	0.393183	0.299864
101	0.3	0.1	0.4	0.1	0.1	0.689582	0.391322	0.297745
102	0.3	0.2	0.1	0.1	0.3	0.688367	0.392625	0.29983
103	0.3	0.2	0.1	0.2	0.2	0.687528	0.394259	0.301489
104	0.3	0.2	0.1	0.3	0.1	0.687632	0.395967	0.301063
105	0.3	0.2	0.2	0.1	0.2	0.689104	0.39207	0.298582
106	0.3	0.2	0.2	0.2	0.1	0.688754	0.394593	0.299162
107	0.3	0.2	0.3	0.1	0.1	0.68986	0.392787	0.296951
108	0.3	0.3	0.1	0.1	0.2	0.68869	0.393212	0.299095
109	0.3	0.3	0.1	0.2	0.1	0.688308	0.395719	0.299722
110	0.3	0.3	0.2	0.1	0.1	0.689902	0.393455	0.296753
111	0.3	0.4	0.1	0.1	0.1	0.689433	0.394242	0.297394
112	0.4	0.1	0.1	0.1	0.3	0.688081	0.391262	0.300388
113	0.4	0.1	0.1	0.2	0.2	0.687231	0.392918	0.30208
114	0.4	0.1	0.1	0.3	0.1	0.687266	0.394546	0.301704
115	0.4	0.1	0.2	0.1	0.2	0.688881	0.390673	0.299014
116	0.4	0.1	0.2	0.2	0.1	0.688476	0.393144	0.299647
117	0.4	0.1	0.3	0.1	0.1	0.689596	0.391248	0.297369
118	0.4	0.2	0.1	0.1	0.2	0.688599	0.392583	0.299084
119	0.4	0.2	0.1	0.2	0.1	0.688203	0.395062	0.299729
120	0.4	0.2	0.2	0.1	0.1	0.68979	0.392795	0.296716
121	0.4	0.3	0.1	0.1	0.1	0.689416	0.393926	0.297151
122	0.5	0.1	0.1	0.1	0.2	0.688244	0.391246	0.299844
123	0.5	0.1	0.1	0.2	0.1	0.687781	0.393666	0.300551
124	0.5	0.1	0.2	0.1	0.1	0.689475	0.3913	0.29734
125	0.5	0.2	0.1	0.1	0.1	0.689237	0.393312	0.297347
126	0.6	0.1	0.1	0.1	0.1	0.688801	0.391825	0.298226

We elaborated an additional sensitivity analysis, reported in the Data S1 (A.2.), starting from a paradoxical scenario that we built to further validate the work's internal validity. We have constructed this scenario to determine how the decision-makers' volition to give high weight to their reference area criteria and low weight to those unrelated to their interests could impact the medical device's choice. Specifically, we considered, on the one hand, that the clinical area evaluators assign maximum importance to the 'medical-sanitary macro-area criteria', and minimum significance to those of the 'sustainability macro-area', and, on the other hand, that the administrative area evaluators do the opposite.

The results obtained, carrying out the sensitivity analysis on this scenario with the same variables (dependent and independent) and the same methodology seen above, fully support the proposed tool's robustness.

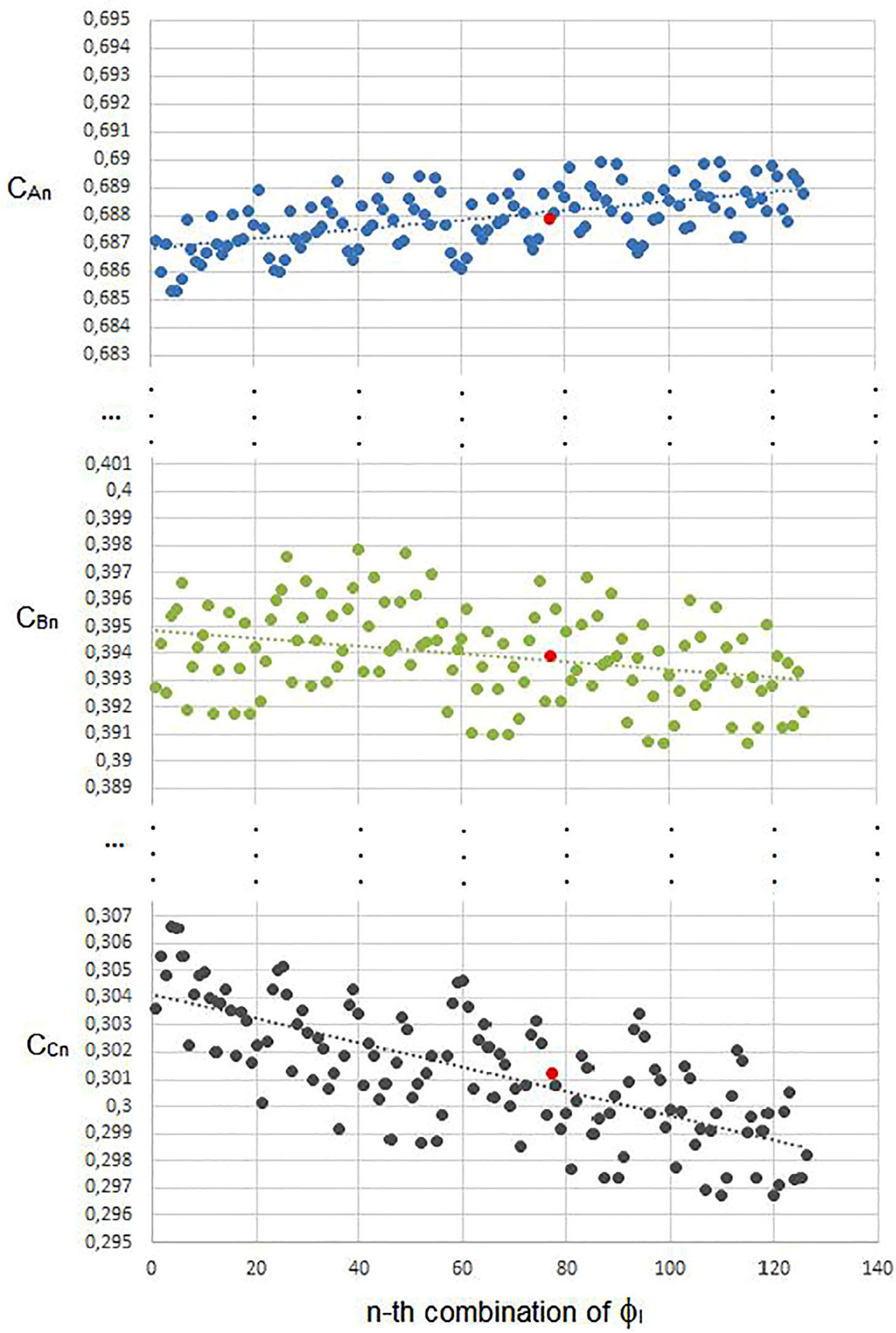

FIGURE 1 Dispersion of the proximity coefficients (C_{in}) as the 'Opinion Value' coefficient (ϕ_l) varies

TABLE 3 Descriptive analysis of the variable 'C_{in}'

Variable	Obs	Mean	SD	Min	Max
C_{An}	126	0.6879065	0.0010716	0.685322	0.689902
C_{Bn}	126	0.3939365	0.0016765	0.3906734	0.3978409
C_{Cn}	126	0.3012781	0.0022858	0.296716	0.306596

5 | DISCUSSION AND CONCLUSION

Because of several reasons (Ciani et al., 2017; Drummond et al., 2008; Dutot et al., 2017; Fuchs et al., 2017; Guerra Bretaña & Flórez-Rendón, 2018; Martelli et al., 2017; Tarricone, Callea, et al., 2017a; Tarricone, Torbica, & Drummond, 2017b), macro-level HTA in many cases failed to provide local institutions/hospitals with guidelines that could help them in assessing medical devices. Therefore, local institutions/hospitals started assessing medical devices locally, many times relying on HTA tools capable to take into account, by means of MCDA, the multi-faceted aspects that new medical devices bring in. However, since this was an emergent process with no consistent development (Martelli et al., 2017), local institutions started using criteria strictly tailored to the specific technology to be evaluated (Sloane, 2004), with obvious flaws in terms of 'portability' of criteria between devices (Morton & Fasolo, 2009).

This article, while proposing a general tool that tackles these critical issues and while giving decision makers at micro-level a methodological guidance able to help them whatever the medical device and the country they are, contributes to the scientific debate that complains about the lack of a consistent development of micro-level HTA assessment (Martelli et al., 2017).

Specifically, this article answers the scientific call which asks for tools capable of taking into account a wide and stable set of criteria (Ivlev et al., 2015; Wahlster et al., 2015), based on clear and above all shared assessments. As a matter of facts, to our best knowledge, this is the first time that an MCDA tool for HTA presents criteria not tailored ad hoc for the specific device which needs to be assessed. Indeed, if on the one hand the state-of-the-art literature already presents general tools – such as for instance ICER or QALY (Marsh et al., 2014; Oliveira et al., 2019; Thokala & Duenas, 2012) – on the other hand, these tools are not multi-criteria, in that they merely investigate only two aspects, that is, the cost and benefits in terms of length of life and health-related quality of life (Devlin & Sussex, 2011; Marsh et al., 2014; Wahlster et al., 2015) associated with the device. This way, important criteria, whose role can be pivotal for the health structure that is assessing a medical device (Guitouni & Martel, 1998), are completely neglected. From this perspective, the tool proposed in this article aims at overtaking this sort of trade-off that contributions in the state-of-the-art literature, although implicitly, assume: tools can be general if the criteria are few, but if a multiplicity of criteria are to be considered, then the tool is specifically tailored to the particular medical device, with no chances of portability.

It is important to note that the tool put forward in this article joins that part of the scientific conversation that calls for tools able to take into account the idiosyncratic characteristics of the context where decisions have to be made (Gagnon, 2014; Sampietro-Colom et al., 2012; Sampietro-Colom & Martin, 2016). Indeed, two are the possibilities for adapting the tool to the local characteristics and necessities. Firstly, by deleting criteria, as well as adding new others able to take into account specific aspects that may be considered as pivotal, this tool can be adapted to any specific situation (Ehlers et al., 2006). Secondly, local decision-makers, by adjusting the criteria weights, can take into account local preferences and priorities. In addition, this tool can be easily adopted in health contexts other than the Italian one. Indeed, it is a very recent news that the Italian Medical Device Classification, explained before and used in this article together with the analysis of the extant literature to identify the criteria, will be the basis for the definition of the European Medical Device Nomenclature that is the nomenclature used by manufacturers when registering their medical devices in the EUDAMED database (European Commission, DG Health and Food Safety, 2020).

But implications are not only at the theoretical level: they are also at the managerial and policy level.

At the managerial level, implications are to be distinguished according to the unit of analysis, whether the local adopting institution/hospital or the manufacturer.

As regards the former, it is valuable to notice that this MCDA tool for HTA, while building on pre-defined criteria and being readily implementable in real contexts whatever the device to be evaluated, allows health institutions to skip time-consuming phases. For instance, this is the case of the design and validation of criteria (Marsh et al., 2016), which vice versa need to be performed anytime new specific criteria are to be conceived. This means that the assessment of devices can be implemented immediately (Blythe et al., 2019; Bretoni et al., 2019). Beneficial consequences are evident. A first benefit can be connected to the speed of introduction of the assessed devices in the health structures for the diagnosis and access to treatment of innumerable pathologies. A second benefit can be connected to the reduction of the opportunity cost: local institutions can reduce the amount of effort devoted to the design of ad-hoc tools. In so doing, resources are released for other activities.

As an addition, the possibility to readily implement the tool and reduce time-consuming phases makes this tool less expensive in terms of man-hour than other tools and hence more suitable also for the assessment of simpler, lower-risk, and cheaper medical devices, which conversely are usually excluded from the application of HTA techniques (Sloane, 2004) because of the effort these techniques require. This is important in that the total cost paid for this type of devices is anyway relevant in the corporate financial statement of local healthcare institutions and basing their assessment only on the criterion of 'lower cost' is rather limited and incomplete. This way, issues connected with limited resources (Barron et al., 2015; Ivlev et al., 2015) and expenditure prioritisation (Devlin & Sussex, 2011) are taken into account.

As an addition, learning economies are also possible: the extended use of the same general tool for different devices nurtures economies connected with the growth of experience.

As regards managerial implications connected to manufacturers, it is worth noting that this article, while proposing a tool which can be used independently of the heterogeneity of devices, aims to cope with a major managerial challenge encountered by manufacturers producing different types of devices and for whom local institutions/hospitals are usually the entry point for their new and innovative medical devices (Dutot et al., 2017). Specifically, this tool would avoid manufacturers to embark in distinct local HTA processes for the different devices they produce. Analogously, a local HTA tool employed at the international level for the assessment of medical devices would allow manufacturers of a single device to devote less resources to cope with different assessment processes brought about by different local institutions in the different countries (Ciani et al., 2017). Indeed, a manufacturer producing a single device intended for commercialization in the international market is challenged by a very large set of heterogeneous local procedures which absorb large manufacturers' resources.

From a policy making point of view, this tool can speed the introduction in local institutions of innovative products which can be beneficial to patients, this way promoting public health. As an addition, a tool like this can be used as a methodological support to initiate device assessment in the low- and middle-income countries that are still lagging behind in the HTA process at the micro-level (Gagnon, 2014) and whose local institutions/hospitals probably do not have enough resources to autonomously build tailored tools.

The main limitation of this work can be identified in the difficulty of finding literature related to tests of experimental medical devices. This is mainly due to the high costs of their trials (Akpinar et al., 2019), unlike what happens with drugs whose tests are multiple, easier to perform and with more available results in the literature. This limitation translates into the difficulty to evaluate those criteria that instead could be better assessed if decision makers could rely on data and analyses from the scientific literature. To overcome this limitation, higher preparation and competence of the decision-makers in the area are necessary.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the Data S1 of this article.

REFERENCES

AAET for Italian Ministry of Health (2016). Medical device data bank, user manual medical device manufacturer profile, version 8.4. https://www.salute.gov.it/imgs/C_17_pagineAree_395_listaFile_itemName_1_file.pdf

Abdel-Basset, M., Manogaran, G., Gamal, A., et al. (2019). A group decision making framework based on neutrosophic TOPSIS approach for smart medical device selection. *Journal of Medical Systems*, *43*, 38. https://doi.org/10.1007/s10916-019-1156-1

Adunlin, G. M. A., Diaby, V., & Xiao, H. (2014). Application of multicriteria decision analysis in health care: A systematic review and bibliometric analysis. *Health Expectations*, *18*, 1894–1905.

AIFA, Agenzia Italiana del Farmaco, (2017). Determina n. 1535/2017, Allegato 1 "Criteri per la valutazione dell'innovatività". https://www.aifa.gov.it/documents/20142/516919/Determina_criteri_classificazione_farmaci_innovativi.pdf/97621d1f-f4f4-01b7-4629-9e8df6f6bf46?t=1542475263114.

Akpinar, I., Ohinmaa, A., Thording, L., Tran, D. T., Fedorak, R. N., Richer, L., & Jacobs, P. (2019). The costs of industry-sponsored medical device clinical trials in Alberta. *PharmacoEconomics Open*, *3*, 591–597.

Aloini, D., Dulmin, R., Mininno, V., Pellegrini, L., & Farina, G. (2017). Technology assessment with IF-TOPSIS: An application in the advanced underwater system sector. *Technological Forecasting and Social Change*, *131*, 38–48.

Aloini, D., Dulmin, R., & Mininno, V. (2014). A peer IF-TOPSIS based decision support system for packaging machine selection. *Expert System with Applications*, *41*(5), 2157–2165.

ASSOBIOMEDICA, (2017). Studi numero 38 - novembre 2017 osservatorio parco installato: le apparecchiature di diagnostica per immagini in Italia. https://www.quotidianosanita.it/allegati/allegato195048.pdf.

Baltussen, R., Marsh, K., Thokala, P., Diaby, V., Castro, H., Cleemput, I., Garau, M., Iskrov, G., Olyaeemanesh, A., Mirelman, A., Mobinizadeh, M., Morton, A., Tringali, M., van Til, J., Valentim, J., Wagner, M., Youngkong, S., Zah, V., Toll, A., … Broekhuizen, H. (2019). Multicriteria decision analysis to support health technology assessment agencies: Benefits, limitations, and the way forward. *Value in Health*, *22*(11), 1283–1288.

Barron, A. J., Klinger, C., Shah, S. M., & Wright, J. S. (2015). A regulatory governance perspective on health technology assessment (HTA) in France: The contextual mediation of common functional pressures. *Health Policy*, *119*(2), 137–146.

Batavia, A. I., & Hammer, G. S. (1990). Toward the development of consumer-based criteria for the evaluation of assistive devices. *Journal of Rehabilitation Research and Development*, *27*(4), 425–436. https://doi.org/10.1682/jrrd.1990.10.0425 PMID: 2089152.

Blythe, R., Naidoo, S., Abbott, C., Bryant, G., Dines, A., & Graves, N. (2019). Development and pilot of a multicriteria decision analysis (MCDA) tool for health services administrators. *BMJ Open*, *9*, e025752.

Bretoni, A., Ferrario, L., & Foglia, E. (2019). HTA and innovative treatments evaluation: The case of metastatic castration-resistant prostate cancer. *ClinicoEconomics and Outcomes Research*, *11*, 515–516.

Chandra, A., & Skinner, J. (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature*, *50*(3), 645–680.

Ciani, O., Britni, W., Van Giessen, A., & Taylor, R. S. (2017). Linking the regulatory and reimbursement processes for medical devices. The need for integrated assessments. *Health Economics*, *26*(Suppl. 1), 13–29.

Devlin, N. J., & Sussex, J. (2011). *Incorporating multiple criteria in HTA: Methods and processes*. Office of Health Economics.

Drummond, M., Schwartz, J., Jönsson, B., Luce, B., Neumann, P., Siebert, U., & Sullivan, S. (2008). Key principles for the improved conduct of health technology assessments for resource allocation decisions. *International Journal of Technology Assessment in Health Care*, *24*(3), 244–258.

Dutot, C., Mercier, G., Borget, I., De Sauvebeuf, C., & Martelli, N. (2017). Hospital-based health technology assessment for the adoption of innovative medical devices within French hospitals: Opportunities and challenges for industry. *International Journal of Technology Assessment in Health Care*, *33*(2), 297–302.

Ehlers, L., Vestergaard, M., Kidholm, K., Bonnevie, B., Pedersen, P. H., Jørgensen, T., Jensen, M. F., Kristensen, F. B., & Kjølby, M. (2006). Doing mini–health technology assessments in hospitals: A new concept of decision support in health care? *International Journal of Technology Assessment in Health Care*, *22*(3), 295–301.

European Commission Enterprise and Industry Directorate General (2009). Guidelines on Medical Device. Clinical Evaluation: a guide for manufactures and notified bodies. https://ec.europa.eu/health/sites/default/files/md_sector/docs/md_emdn_eudamed_nomenclature_en.pdf

European Commission, DG Health and Food Safety (2020). The European Medical Device Nomenclature (EMDN). https://ec.europa.eu/docsroom/documents/10336/attachments/1/translations/en/renditions/native

Favaretti, C., Cicchetti, A., Guarrera, G., Marchetti, M., & Ricciardi, W. (2009). Health technology assessment in Italy. *International Journal of Technology Assessment in Health Care*, *25*(Suppl. 1), 127–133.

FDA, U.S. Food & Drug Administration Medical Device Databases n.d. www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/medical-device-databases.

Fuchs, S., Olberg, B., Panteli, D., Perleth, M., & Busse, R. (2017). HTA of medical devices: Challenges and ideas for the future from a European perspective. *Health Policy*, *121*(3), 215–229.

Gagnon, M. P. (2014). Hospital –based health technology assessment: Developments to date. *PharmacoEconomics*, *32*(9), 819–824.

Gagnon, M. P., Desmartis, M., Poder, T., & Witteman, W. (2014). Effects and repercussions of local/hospital-based health technology assessment (HTA): A systematic review. *Systematic Reviews*, *3*(129), 1–14.

Guerra Bretaña, R., & Flórez-Rendón, A. (2018). Impact of regulations on innovation in the field of medical devices. *Research on Biomedical Engineering*, *34*, 356–367.

Guitouni, A., & Martel, J. M. (1998). Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, *109*(2), 501–521.

Ivlev, I., Vacek, J., & Kneppo, P. (2015). Multi-criteria decision analysis for supporting the selection of medical devices under uncertainty. *European Journal of Operational Research*, *247*(1), 216–228.

Lazzini S. (2014). Strumenti economico-aziendali per il governo della sanità. Franco Angeli ed., EAN: 9788820462871, ISBN: 8820462877.

Marsh, K., IJzerman, M., Thokala, P., Baltussen, R., Boysen, M., Kaló, Z., Lönngren, T., Mussen, F., Peacock, S., Watkins, J., & Devlin, N. (2016). Multiple criteria decision analysis for health care decision making-emerging good practices: Report 2 of the ISPOR MCDA emerging good practices task force. *Value in Health*, *19*(2), 125–137.

Marsh, K., Lanitis, T., Neasham, D., Orfanos, P., & Caro, J. (2014). Assessing the value of healthcare interventions using multi-criteria decision analysis: A review of the literature. *PharmacoEconomics*, *32*(4), 345–365.

Marsh, K., Zaiser, E., Orfanos, P., Salverda, S., Wilcox, T., Sun, S., & Dixit, S. (2017). Evaluation of COPD treatments: a multicriteria decision analysis of aclidinium and tiotropium in the united states. *Value in Health*, *20*, 132–140. https://doi.org/10.1016/j.jval.2016.08.724.

Martelli, N., Hansen, P., van den Brink, H., Boudard, A., Cordonnierd, A. L., Devaux, C., Pineau, J., Prognon, P., & Borget, I. (2016). Combining multi-criteria decision analysis and mini-health technology assessment: A funding decision-support tool for medical devices in a university hospital setting. *Journal of Biomedical Informatics*, *59*(February), 201–208.

Martelli, N., Devaux, C., Van den Brink, H., Billaux, M., Pineau, J., Prognon, P., & Borget, I. (2017). Harmonizing health technology assessment practices in university hospitals: To what extent is the mini-HTA model suitable in the French context? *International Journal of Technology Assessment in Health Care*, *33*(2), 307–314.

Mathew, J. (2011). KNOW ESSENTIALS: A tool for informed decisions in the absence of formal HTA systems. *International Journal of Technology Assessment in Health Care*, *27*(2), 139–150.

Miller, L. J., & Wei, L. (2018). *These are the economies with the Most (and least) efficient health care*. [online]. Bloomberg Accessed June 19, 2021. https://www.bloomberg.com/news/articles/2018-09-19/u-s-near-bottom-of-health-index-hong-kong-and-singapore-at-top

Morton, A., & Fasolo, B. (2009). Behavioural decision theory for multi-criteria decision analysis: A guided tour. *Journal of Operational Research Society*, *60*, 268–275.

Oliveira, M. D., Mataloto, I., & Kanavos, P. (2019). Multi-criteria decision analysis for health technology assessment: Addressing methodological challenges to improve the state of the art. *The European Journal of Health Economics*, *20*, 891–918.

Ottardi, C., Damonti, A., Porazzi, E., Foglia, E., Ferrario, L., Villa, T., Aimar, E., Brayda-Bruno, M., & Galbusera, F. (2017). A comparative analysis of a disposable and a reusable pedicle screw instrument kit for lumbar arthrodesis: Integrating HTA and MCDA. *Health Economic Review*, *7*(17).

Park, J., Kim, E., & Shin, K. (2019). Developing an evaluation framework for selecting optimal medical devices. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(3), 64.

Pecchia, L., Pallikarakis, N., Magjarevic, R., & Iadanzade, E. (2019). Health technology assessment and biomedical engineering: Global trends, gaps and opportunities. *Medical Engineering & Physics*, *72*, 19–26.

Sampietro-Colom, L., Morilla-Bachs, I., Gutierrez-Moreno, S., & Gallo, P. (2012). Development and test of a decision support tool for hospital health technology assessment. *International Journal of Technology Assessment in Health Care*, *28*(4), 460–465.

Sampietro-Colom, L., & Martin, J. (2016), Hospital-Based Health Technology Assessment. Adis (Eds.).

Sloane, E. B. (2004). Using a decision support system tool for healthcare technology assessments. *IEEE Engineering in Medicine and Biology Magazine*, *23*, 42–55.

Tarricone, R., Callea, G., Ogorecv, M., & Pevolnik, R. V. (2017a). Improving the methods for the economic evaluation of medical devices. *Health Economics*, *26*(Suppl. 1), 70–92.

Tarricone, R., Torbica, A., & Drummond, M. (2017b). Challenges in the assessment of medical devices: The MEDTECHTA project. *Health Economics*, *26*(Suppl. 1), 5–12.

Thokala, P., & Duenas, A. (2012). Multiple criteria decision analysis for health technology assessment. *Value in Health*, *15*, 1172–1178.

Thokala, P., Devlin, N., Marsh, K., Baltussen, R., Boysen, M., Kalo, Z., Longrenn, T., Mussen, F., Peacock, S., Watkins, J., Pharm, D., & Ijzerman, M. (2016). Multiple criteria decision analysis for health care decision making-an introduction: Report 1 of the ISPOR MCDA emerging good practices task force. *Value in Health*, *19*(1), 1–13.

Vettoretto, N., Foglia, E., Ferrario, L., Arezzo, A., Cirocchi, R., Cocorullo, G., Currò, G., Marchi, D., Portale, G., Gerardi, C., Nocco, U., Tringali, M., Anania, G., Piccoli, M., Silecchia, G., Morino, M., Valeri, A., & Lettieri, E. (2018). Why laparoscopists may opt for three-dimensional view: A summary of the full HTA report on 3D versus 2D laparoscopy by S.I.C.E. (Società Italiana di Chirurgia Endoscopica e Nuove Tecnologie). *Surgical Endoscopy*, *32*, 2986–2993.

Wahlster, P., Goetghebeur, M., Kriza, C., Niederländer, C., & Kolominsky-Rabas, B. (2015). Balancing costs and benefits at different stages of medical innovation: A systematic review of multi-criteria decision analysis (MCDA). *BMC Health Services Research*, *15*, 262.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Tallarico, S., Aloini, D., Dulmin, R., Lazzini, S., Mininno, V., & Pellegrini, L. (2021). Health Technology Assessment of medical devices. Overcoming the critical issues of current assessment. *Journal of Multi-Criteria Decision Analysis*, 1–23. https://doi.org/10.1002/mcda.1764

APPENDIX A.: IF-TOPSIS PEER-TO-PEER IMPLEMENTATION

In the following, we report the theoretical details about the eight steps of the peer-to-peer IF-TOPSIS method implementation.

Step I. Construction of the Aggregated Importance IF decision matrix. All the decision-makers vote the importance of the others (Tables A1 and A3) according to an Intuitionistic Fuzzy scale, assuming i_k as the Intuitionistic Fuzzy Number (IFN) associated with this importance.

The opinion of each decision-maker (l is the number of decision makers involved in the decision-making process, in this case, three) has the same value and is represented by the coefficient ϕ (which is called "Opinion Value") defined in the following equation:

$$\phi_1 = \phi_2 = \ldots = \phi_l = \frac{1}{l} = \frac{1}{3}.$$

In the decision-making group, all individual opinions must be grouped in a group opinion. To do this, the IFWA (Intuitionistic Fuzzy Weighted Averaging) operator was used:

$$D_k = \phi_1 i_1{}^{(1)} \phi_2 i_2{}^{(2)} \ldots \phi_l i_l{}^{(l)} = \left[1 - \prod_{k=1}^{l} (1-\mu_k)^{\phi},\ \prod_{k=1}^{l} (v_k)^{\phi},\ \prod_{k=1}^{l} (1-\mu_k)^{\phi} - \prod_{k=1}^{l} (v_k)^{\phi}\right],$$

where: $D_k = [\mu_k, v_k, \pi_k]$ is the IFN which represents the aggregate importance of the k-th decision-maker (Table A3).

The linguistic terms used for the evaluation of the importance of the decision-makers (and also for the criteria) are reported in the following table (Table A2).

Step II. Calculation of the weight of each decision-maker.

TABLE A1 Valuation of decision makers importance (linguistic terms)

	DM 1	DM 2	DM 3	DM4	DM5
DM 1	M	I	VI	I	VI
DM 2	I	I	VI	I	VI
DM 3	M	VI	VI	I	VI
DM4	I	I	VI	VI	VI
DM5	M	I	VI	I	VI

TABLE A2 Translation of linguistic terms in IFN

	Intuitionistic fuzzy numbers		
Linguistic terms	μ	ν	π
Very Important (VI)	0.9	0.05	0.05
Important (I)	0.65	0.25	0.1
Medium (M)	0.5	0.4	0.1
Unimportant (UI)	0.35	0.55	0.1
Very Unimportant (UVI)	0.15	0.8	0.05

TABLE A3 1 - valuation of DM importance (IFN); 2 - aggregate importance of the k-th decision-maker (IFN); 3 - weights of the k-th decision-maker

	DM 1			DM 2			DM 3			DM4			DM5		
	μ	ν	π	μ	ν	π	μ	ν	π	μ	ν	π	μ	ν	π
[1] DM 1	0.5	0.4	0.1	0.65	0.25	0.1	0.9	0.05	0.05	0.65	0.25	0.1	0.9	0.05	0.05
[1] DM 2	0.65	0.25	0.1	0.65	0.25	0.1	0.9	0.05	0.05	0.65	0.25	0.1	0.9	0.05	0.05
[1] DM 3	0.5	0.4	0.1	0.9	0.05	0.05	0.9	0.05	0.05	0.65	0.25	0.1	0.9	0.05	0.05
[1] DM 4	0.65	0.25	0.1	0.65	0.25	0.1	0.9	0.05	0.05	0.9	0.05	0.05	0.9	0.05	0.05
[1] DM 5	0.5	0.4	0.1	0.65	0.25	0.1	0.9	0.05	0.05	0.65	0.25	0.1	0.9	0.05	0.05
[2] Dk	0.56648	0.331445	0.102075	0.72757	0.181195	0.091235	0.9	0.05	0.05	0.72757	0.181195	0.091235	0.9	0.05	0.05
[3] λ_k	0.152871491			0.19400167			0.229562584			0.19400167			0.229562584		

The judgment of decision-makers is expressed in terms of literary judgments which are then transposed into IFNs.

The weights λ_k of the k-th decision-maker (Table A3) can be obtained with the method:

$$\lambda_k = \frac{\left(\mu_k + \pi_k * \left(\frac{\mu_k}{\mu_k + v_k}\right)\right.}{\sum_{k=1}^{l}\left(\mu_k + \pi_k * \left(\frac{\mu_k}{\mu_k + v_k}\right),\right.}$$

$$\sum_{k=1}^{l} \lambda_k = 1.$$

Step III. Construction of the aggregated IF decision matrix.

$$R^{(k)} = \left(r_{ij}{}^{(k)}\right)_{m \times n},$$

where R is the IF decision-making matrix of each decision-maker.

$$R = \left(r_{ij}\right)_{m \times n},$$

where:

$$r_{ij} = IFWA_{\lambda}\left(r_{ij}{}^{(1)}, r_{ij}{}^{(2)}, \ldots, r_{ij}{}^{(l)}\right) = \lambda_1 r_{ij}{}^{(1)} \lambda_2 r_{ij}{}^{(2)} \ldots \lambda_l r_{ij}{}^{(l)}$$
$$= \left[1 - \prod_{k=1}^{l}\left(1 - \mu_{ij}{}^{(k)}\right)^{\lambda_k}, \prod_{k=1}^{l}\left(v_{ij}{}^{(k)}\right)^{\lambda_k}, \prod_{k=1}^{l}\left(1 - \mu_{ij}{}^{(k)}\right)^{\lambda_k} - \prod_{k=1}^{l}\left(v_{ij}{}^{(k)}\right)^{\lambda_k}\right].$$

$$R' = \begin{bmatrix} \left(\mu_{A_1W}(x_1), v_{A_1W}(x_1), \pi_{A_1W}(x_1)\right) & \left(\mu_{A_1W}(x_2), v_{A_1W}(x_2), \pi_{A_1W}(x_2)\right) \ldots & \left(\mu_{A_1W}(x_n), v_{A_1W}(x_n), \pi_{A_1W}(x_n)\right) \\ \left(\mu_{A_2W}(x_1), v_{A_2W}(x_1), \pi_{A_2W}(x_1)\right) \vdots & \left(\mu_{A_2W}(x_2), v_{A_2W}(x_2), \pi_{A_2W}(x_2)\right) \vdots \ddots & \left(\mu_{A_2W}(x_n), v_{A_2W}(x_n), \pi_{A_2W}(x_n)\right) \vdots \\ \left(\mu_{A_mW}(x_1), v_{A_mW}(x_1), \pi_{A_mW}(x_1)\right) & \left(\mu_{A_mW}(x_2), v_{A_mW}(x_2), \pi_{A_mW}(x_2)\right) \ldots & \left(\mu_{A_mW}(x_n), v_{A_mW}(x_n), \pi_{A_mW}(x_n)\right) \end{bmatrix}.$$

and

$$r_{ij} = \left(\mu_{A_i}(x_j), v_{A_i}(x_j), \pi_{A_i}(x_j)\right),$$

$$(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n).$$

The aggregated decision matrix IF can be defined as follows:

$$R = \begin{bmatrix} \left(\mu_{A_1}(x_1), v_{A_1}(x_1), \pi_{A_1}(x_1)\right) & \left(\mu_{A_1}(x_2), v_{A_1}(x_2), \pi_{A_1}(x_2)\right) \ldots & \left(\mu_{A_1}(x_n), v_{A_1}(x_n), \pi_{A_1}(x_n)\right) \\ \left(\mu_{A_2}(x_1), v_{A_2}(x_1), \pi_{A_2}(x_1)\right) \vdots & \left(\mu_{A_2}(x_2), v_{A_2}(x_2), \pi_{A_2}(x_2)\right) \vdots \ddots & \left(\mu_{A_2}(x_n), v_{A_2}(x_n), \pi_{A_2}(x_n)\right) \vdots \\ \left(\mu_{A_m}(x_1), v_{A_m}(x_1), \pi_{A_m}(x_1)\right) & \left(\mu_{A_m}(x_2), v_{A_m}(x_2), \pi_{A_m}(x_2)\right) \ldots & \left(\mu_{A_m}(x_n), v_{A_m}(x_n), \pi_{A_m}(x_n)\right) \end{bmatrix}.$$

Step IV. Calculation of the weights of the individual criteria.
The criteria cannot be assumed of the same value, so we assume that W represents a set of degrees of importance.

To obtain W (Table A4), the individual opinions of each decision-maker regarding the importance of the various criteria must be combined in an overall judgment, as follows:

$$w_j^{(k)} = \left[\mu_j^{(k)}, v_j^{(k)}, \pi_j^{(k)}\right].$$

The criteria weights were calculated using the IFWA operator:

$$w_j = IFWA_{\lambda}\left(w_j{}^{(1)}, w_j{}^{(2)}, \ldots, w_j{}^{(l)}\right) = \lambda_1 w_j{}^{(1)} \lambda_2 w_j{}^{(2)} \ldots \lambda_l w_j{}^{(l)}$$
$$= \left[1 - \prod_{k=1}^{l}\left(1 - \mu_j{}^{(k)}\right)^{\lambda_k}, \prod_{k=1}^{l}\left(v_j{}^{(k)}\right)^{\lambda_k}, \prod_{k=1}^{l}\left(1 - \mu_j{}^{(k)}\right)^{\lambda_k} - \prod_{k=1}^{l}\left(v_j{}^{(k)}\right)^{\lambda_k}\right],$$

$$W = \left[w_1, w_2, \ldots, w_j\right]$$

In the formula above $w_j = \left(\mu_j, v_j, \pi_j\right)$ and $(j = 1, 2, \ldots, n)$.

Step V. Construction of the Aggregated Weighted IF Decision Matrix. After having built the aggregate decision matrix (Step III) and established the weight W of the criteria (Step IV), the weighted aggregate decision IF matrix was constructed:

$$R \otimes W = \{x, \mu_{Ai}(x) * \mu_W(x), v_{Ai}(x) + \; v_W(x) - v_{Ai}(x) * v_W(x) | x \in X\},$$

and

$$\pi_{Ai*W}(x) = 1 - v_{Ai}(x) - v_W(x) - \mu_{Ai}(x) * \mu_W(x) + v_{Ai}(x) * v_W(x).$$

In this way it is possible to define the aggregate decision weighted IF matrix as follows:

Step VI. Calculation of the intuitionistic fuzzy positive-ideal solution and the intuitionistic fuzzy negative-ideal solution;
J1 and J2 have been defined as two sets that contain benefit and cost criteria, respectively.

The criteria of benefit contribute to obtaining the positive – ideal solutions, while those of cost contribute to the negative – ideal solutions. The components of A+ and A− as can be seen from the formulas below, are respectively maximized/minimized and minimized/maximized. This allows choosing the alternatives that have at the same time the minimum distance from positive solutions – ideal A+ and the maximum from negative ones – ideal A− (Table A4).

TABLE A4 1 - weights of the individual criteria; 2 - A+ the positive - ideal IF solution; 3 - A− the negative - ideal IF solution

Criteria	[1] W			[2] A+			[3] A−		
	μ	ν	π	μ	ν	π	μ	ν	π
Dimensions adequacy	0.802276	0.120639	0.077085	0.67838	0.202663	0.118954044	0.46375252	0.400232	0.136015595
Weight adequacy	0.680751	0.223253	0.095996	0.61268	0.26209	0.125233651	0.422024768	0.439146	0.138829529
Life-time adequacy	0.552516	0.344147	0.103337	0.33366	0.536391	0.129944838	0.333664283	0.536391	0.129944838
Number of detected parameters	0.5	0.4	0.1	0.33454	0.539608	0.125849366	0.075	0.88	0.045
Power supply type compatibility	0.624924	0.18927	0.185806	0.56243	0.229807	0.207761706	0.406200596	0.391953	0.201846746
Battery recharge time compatibility	0.815495	0.110125	0.074379	0.73395	0.154619	0.111435228	0.326198114	0.555063	0.118739246
Electric consumption	0.747122	0.165784	0.087094	0.29885	0.582892	0.118259116	0.298848742	0.582892	0.118259116
Alarm accuracy	0.737475	0.172776	0.089749	0.58564	0.277369	0.136989182	0.338922138	0.534346	0.126731685
Number of alarms	0.563755	0.334193	0.102052	0.2255	0.667097	0.107401322	0.366440911	0.500645	0.132914221
Effectiveness level	0.872488	0.068324	0.059188	0.50434	0.36455	0.131111287	0.785239451	0.114908	0.099853038
Efficiency level	0.9	0.05	0.05	0.36	0.525	0.115	0.785239451	0.114908	0.099853038
Satisfaction	0.592921	0.305069	0.10201	0.27249	0.608816	0.118695372	0.533628665	0.339815	0.126555879
Versatility with respect to the types of patients	0.9	0.05	0.05	0.135	0.81	0.055	0.785239451	0.114908	0.099853038
Reduction in drug consumption	0.878892	0.063947	0.057161	0.131834	0.812789	0.055376715	0.5373721	0.273256	0.189371489
Disability reduction	0.9	0.05	0.05	0.36	0.525	0.115	0.585	0.2875	0.1275
Fulfilment of a therapeutic need	0.872488	0.068324	0.059188	0.40992	0.465634	0.124444144	0.761235796	0.149736	0.089028402
Added therapeutic/ technological value	0.83	0.098862	0.071138	0.38254	0.493501	0.123961968	0.747000404	0.143919	0.109080843
Software break risk	0.452362	0.446718	0.10092	0.15921	0.756085	0.084708162	0.276582874	0.570436	0.152980702
Hardware break risk	0.495415	0.401028	0.103556	0.25354	0.630777	0.115687769	0.432243917	0.441952	0.12580369
Device cost	0.736773	0.174055	0.089172	0.19642	0.715536	0.088048386	0.384940869	0.390598	0.224461338
Cost of the diagnostic/ surgical procedure	0.838538	0.092529	0.068932	0.25713	0.648381	0.094486791	0.469781206	0.430826	0.099393242
Trial cost	0.736773	0.174055	0.089172	0.11052	0.834811	0.054673058	0.611522146	0.278487	0.109991348
Patient's Indirect costs	0.478985	0.417062	0.103953	0.127692	0.79923	0.073077653	0.292860771	0.547412	0.159727189
Cost of engineering support	0.677147	0.226836	0.096017	0.145802	0.782105	0.07209309	0.590802726	0.294397	0.114800484
Battery sustainability	0.646829	0.252999	0.100173	0.420439	0.439749	0.139812303	0.420438712	0.399017	0.18054383

In the analysed case:

J1 = {Dimensions adequacy, Weight adequacy, Length of lifetime, Number of detected parameters, Power supply type compatibility, Battery recharge time compatibility, Alarm accuracy, Number of alarms, Effectiveness level, Efficiency level, Satisfaction, Versatility with respect to the types of patients, Reduction in drug consumption, Disability reduction, Fulfilment of a therapeutic need, Added therapeutic/technological value}

J2 = {Electric consumption, Software break risk, Hardware break risk, Device cost, Cost of the diagnostic/surgical procedure, Trial cost, Patient's indirect costs}

Now, let us denote by A+ the positive – ideal IF solution, while with A− the negative – ideal IF solution. A+ and A− have been so obtained:

$$A^{+} = \left(\mu_{A^{+}W}(x_j), v_{A^{+}W}(x_j)\right),$$

$$A^{-} = \left(\mu_{A^{-}W}(x_j), v_{A^{-}W}(x_j)\right),$$

where:

$$\mu_{A^{+}W}(x_j) = \left(\left(\max_i \mu_{A_iW}(x_j) | j \in J_1\right) \left(\min_i \mu_{A_iW}(x_j) | j \in J_2\right)\right),$$

$$v_{A^{+}W}(x_j) = \left(\left(\min_i v_{A_iW}(x_j) | j \in J_1\right) \left(\max_i v_{A_iW}(x_j) | j \in J_2\right)\right),$$

$$\mu_{A^{-}W}(x_j) = \left(\left(\min_i \mu_{A_iW}(x_j) | j \in J_1\right) \left(\max_i \mu_{A_iW}(x_j) | j \in J_2\right)\right),$$

$$v_{A^{-}W}(x_j) = \left(\left(\max_i v_{A_iW}(x_j) | j \in J_1\right) \left(\min_i v_{A_iW}(x_j) | j \in J_2\right)\right).$$

Step VII. Calculation of the separation measures and the proximity coefficients.

Once the flows related to the solutions are obtained, it is important to measure the distance between the various alternatives on the Intuitionistic Fuzzy Set.

The separation measurements, indicated with S+ and S−, of any alternative technology with positive fuzzy intuitionistic solutions – ideal and negative – ideal are calculated to make the final choice.

$$S^{+}=\sqrt{\frac{1}{2n}\sum_{j=1}^{n}\left[\left(\mu_{A_iW}(x_j)-\mu_{A^{+}W}(x_j)\right)^2+\left(v_{A_iW}(x_j)-v_{A^{+}W}(x_j)\right)^2+\left(\pi_{A_iW}(x_j)-\pi_{A^{+}W}(x_j)\right)^2\right]},$$

$$S^{-}=\sqrt{\frac{1}{2n}\sum_{j=1}^{n}\left[\left(\mu_{A_iW}(x_j)-\mu_{A^{-}W}(x_j)\right)^2+\left(v_{A_iW}(x_j)-v_{A^{-}W}(x_j)\right)^2+\left(\pi_{A_iW}(x_j)-\pi_{A^{-}W}(x_j)\right)^2\right]}.$$

The relative closeness coefficient concerning the ideal solution is:

$$C_i=\frac{S_{i^-}}{S_{i^-}+S_{i^+}},$$

where $0 \le C_i \le 1$.

For each device the obtained results are shown in Table A5.

TABLE A5 Separation measurements and relative closeness coefficient

	a	b	c
S^+	0.287057	0.145959826	0.124195
S^-	0.130209	0.224577	0.288075
C_i	0.687948	0.393914691	0.301247

Step VIII. Rank of alternatives according to the decreasing proximity coefficient.

After the relative closeness coefficient of each alternative is determined, alternatives are ranked according to the descending order of C_i. The final ranking is the following: A – B – C.

TABLE A6 Criteria importance evaluation (paradoxical scenario)

Criteria	DM 1	DM 2	DM 3	DM 4	DM 5
Dimensions adequacy	VI	VI	VI	VU	VU
Weight adequacy	VI	VI	I	VU	VU
Life-time adequacy	VI	VI	U	VU	VU
Number of detected parameters	VI	VI	M	VU	VU
Power supply type compatibility	VI	VI	I	VU	VU
Battery recharge time compatibility	VI	VI	VI	VU	VU
Electric consumption	VU	VU	U	VI	VI
Alarm accuracy	VI	VI	I	VU	VU
Number of alarms	VI	VI	I	VU	VU
Effectiveness level	VI	VI	VI	VU	VU
Efficiency level	VI	VI	VI	VU	VU
Satisfaction	VI	VI	I	VU	VU
Versatility with respect to the types of patients	VI	VI	VI	VU	VU
Reduction in drug consumption	VI	VI	I	VU	VU
Disability reduction	VI	VI	VI	VU	VU
Fulfilment of a therapeutic need	VI	VI	VI	VU	VU
Added therapeutic/technological value	VI	VI	VI	VU	VU
Software break risk	VU	VU	U	VI	VI
Hardware break risk	VU	VU	U	VI	VI
Device cost	VU	VU	U	VI	VI
Cost of the diagnostic/surgical procedure	VU	VU	I	VI	VI
Trial cost	VU	VU	U	VI	VI
Patient's Indirect costs	VU	VU	VU	VI	VI
Cost of engineering support	VU	VU	VU	VI	VI
Battery sustainability	VU	VU	U	VU	VU

APPENDIX B.: PARADOXICAL SCENARIO

This scenario includes the following assumptions for the criteria importance evaluation (Table A6):

1. The physicians (DM1 and DM2) evaluate the medical criteria as VI while the economic-managerial criteria as VU;

2. The senior administrative managers (DM4 and DM5) evaluate the economic-managerial criteria as VI while the medical criteria as VU;

3. The importance criteria assessment made by DM3 (physician and manager) remain unchanged.

The method was implemented with the 126 possible combinations related to the value of the decision-makers' opinions (Table A7); the ranking of the alternatives remains unchanged.

The following figure (Figure A1) shows the proximity coefficient dispersion obtained in the 126 different combinations of the DMs opinion values for each of the three alternatives.

TABLE A7 Variation of the proximity coefficients C_{in} as the 'Opinion Value' coefficients vary (paradoxical scenario)

	'Opinion Value' coefficients ϕ_i					Proximity coefficient C_{in}		
n-combinations	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
1	0.1	0.1	0.1	0.1	0.6	0.676495	0.423511	0.314615
2	0.1	0.1	0.1	0.2	0.5	0.676943	0.428027	0.314887
3	0.1	0.1	0.1	0.3	0.4	0.677251	0.429961	0.314846
4	0.1	0.1	0.1	0.4	0.3	0.677322	0.430746	0.314864
5	0.1	0.1	0.1	0.5	0.2	0.676998	0.431038	0.315147
6	0.1	0.1	0.1	0.6	0.1	0.675193	0.431269	0.316553
7	0.1	0.1	0.2	0.1	0.5	0.678192	0.421361	0.312294
8	0.1	0.1	0.2	0.2	0.4	0.678662	0.427096	0.31251
9	0.1	0.1	0.2	0.3	0.3	0.678904	0.427563	0.312515
10	0.1	0.1	0.2	0.4	0.2	0.678702	0.428319	0.312739
11	0.1	0.1	0.2	0.5	0.1	0.677128	0.428853	0.313956
12	0.1	0.1	0.3	0.1	0.4	0.678502	0.420961	0.311862
13	0.1	0.1	0.3	0.2	0.3	0.678922	0.425269	0.312106
14	0.1	0.1	0.3	0.3	0.2	0.678896	0.427104	0.312312
15	0.1	0.1	0.3	0.4	0.1	0.677441	0.42809	0.313474
16	0.1	0.1	0.4	0.1	0.3	0.678499	0.420853	0.311818
17	0.1	0.1	0.4	0.2	0.2	0.678655	0.421389	0.31226
18	0.1	0.1	0.4	0.3	0.1	0.677374	0.427181	0.3134
19	0.1	0.1	0.5	0.1	0.2	0.67818	0.420719	0.3103
20	0.1	0.1	0.5	0.2	0.1	0.677075	0.425152	0.313401
21	0.1	0.1	0.6	0.1	0.1	0.676592	0.420358	0.313111
22	0.1	0.2	0.1	0.1	0.5	0.672184	0.423835	0.318564
23	0.1	0.2	0.1	0.2	0.4	0.672142	0.427893	0.319338
24	0.1	0.2	0.1	0.3	0.3	0.672193	0.42961	0.319547
25	0.1	0.2	0.1	0.4	0.2	0.671932	0.430304	0.319833
26	0.1	0.2	0.1	0.5	0.1	0.670397	0.430762	0.321
27	0.1	0.2	0.2	0.1	0.4	0.673832	0.421905	0.316362
28	0.1	0.2	0.2	0.2	0.3	0.673843	0.425855	0.317037
29	0.1	0.2	0.2	0.3	0.2	0.673682	0.42753	0.317386
30	0.1	0.2	0.2	0.4	0.1	0.67232	0.428452	0.318457
31	0.1	0.2	0.3	0.1	0.3	0.674092	0.421522	0.315976
32	0.1	0.2	0.3	0.2	0.2	0.67389	0.425447	0.316791
33	0.1	0.2	0.3	0.3	0.1	0.672622	0.42733	0.317928
34	0.1	0.2	0.4	0.1	0.2	0.673869	0.42135	0.316076
35	0.1	0.2	0.4	0.2	0.1	0.672554	0.425429	0.317674
36	0.1	0.2	0.5	0.1	0.1	0.672476	0.421012	0.316961

TABLE A7 (Continued)

n-combinations	'Opinion Value' coefficients ϕ_i					Proximity coefficient C_{in}		
	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
37	0.1	0.3	0.1	0.1	0.4	0.670802	0.423934	0.319791
38	0.1	0.3	0.1	0.2	0.3	0.670547	0.427843	0.320766
39	0.1	0.3	0.1	0.3	0.2	0.670288	0.429485	0.321221
40	0.1	0.3	0.1	0.4	0.1	0.668858	0.43033	0.322349
41	0.1	0.3	0.2	0.1	0.3	0.672351	0.422064	0.317698
42	0.1	0.3	0.2	0.2	0.2	0.671999	0.425884	0.318674
43	0.1	0.3	0.2	0.3	0.1	0.670714	0.427702	0.319829
44	0.1	0.3	0.3	0.1	0.2	0.672389	0.421622	0.317459
45	0.1	0.3	0.3	0.2	0.1	0.670996	0.425577	0.319142
46	0.1	0.3	0.4	0.1	0.1	0.671114	0.421226	0.318214
47	0.1	0.4	0.1	0.1	0.3	0.67024	0.423945	0.320269
48	0.1	0.4	0.1	0.2	0.2	0.669689	0.427784	0.321473
49	0.1	0.4	0.1	0.3	0.1	0.668279	0.429552	0.322746
50	0.1	0.4	0.2	0.1	0.2	0.671552	0.422032	0.318344
51	0.1	0.4	0.2	0.2	0.1	0.670092	0.425944	0.320099
52	0.1	0.4	0.3	0.1	0.1	0.670545	0.421364	0.318754
53	0.1	0.5	0.1	0.1	0.2	0.669759	0.423839	0.320616
54	0.1	0.5	0.1	0.2	0.1	0.668064	0.427744	0.322618
55	0.1	0.5	0.2	0.1	0.1	0.670026	0.421709	0.319342
56	0.1	0.6	0.1	0.1	0.1	0.668309	0.4234	0.321515
57	0.2	0.1	0.1	0.1	0.5	0.672531	0.420839	0.317636
58	0.2	0.1	0.1	0.2	0.4	0.672476	0.424873	0.318497
59	0.2	0.1	0.1	0.3	0.3	0.672515	0.42659	0.318741
60	0.2	0.1	0.1	0.4	0.2	0.672242	0.427265	0.319036
61	0.2	0.1	0.1	0.5	0.1	0.670677	0.427607	0.320175
62	0.2	0.1	0.2	0.1	0.4	0.674347	0.419072	0.315405
63	0.2	0.1	0.2	0.2	0.3	0.674331	0.422995	0.316163
64	0.2	0.1	0.2	0.3	0.2	0.674152	0.424653	0.31654
65	0.2	0.1	0.2	0.4	0.1	0.672769	0.425476	0.317589
66	0.2	0.1	0.3	0.1	0.3	0.674641	0.418719	0.315012
67	0.2	0.1	0.3	0.2	0.2	0.674408	0.4226	0.315904
68	0.2	0.1	0.3	0.3	0.1	0.673112	0.424389	0.317038
69	0.2	0.1	0.4	0.1	0.2	0.674423	0.418537	0.315101
70	0.2	0.1	0.4	0.2	0.1	0.673067	0.422497	0.316744
71	0.2	0.1	0.5	0.1	0.1	0.673013	0.418129	0.315936
72	0.2	0.2	0.1	0.1	0.4	0.669594	0.421432	0.320252
73	0.2	0.2	0.1	0.2	0.3	0.669158	0.425144	0.321481
74	0.2	0.2	0.1	0.3	0.2	0.66883	0.426696	0.322021
75	0.2	0.2	0.1	0.4	0.1	0.667449	0.427427	0.323056
76	0.2	0.2	0.2	0.1	0.3	0.671216	0.419774	0.318245
77	0.2	0.2	0.2	0.2	0.2	0.670703	0.423407	0.319433
78	0.2	0.2	0.2	0.3	0.1	0.66941	0.425068	0.320574
79	0.2	0.2	0.3	0.1	0.2	0.671286	0.419364	0.317998
80	0.2	0.2	0.3	0.2	0.1	0.669789	0.423063	0.319799
81	0.2	0.2	0.4	0.1	0.1	0.670073	0.418916	0.31864

(Continues)

TABLE A7 (Continued)

	'Opinion Value' coefficients ϕ_i					Proximity coefficient C_{in}		
n-combinations	DM1	DM2	DM3	DM4	DM5	C_{An}	C_{Bn}	C_{Cn}
82	0.2	0.3	0.1	0.1	0.3	0.668579	0.421603	0.321111
83	0.2	0.3	0.1	0.2	0.2	0.667819	0.425189	0.322587
84	0.2	0.3	0.1	0.3	0.1	0.66642	0.426776	0.323826
85	0.2	0.3	0.2	0.1	0.2	0.669931	0.419914	0.319306
86	0.2	0.3	0.2	0.2	0.1	0.668348	0.42352	0.321201
87	0.2	0.3	0.3	0.1	0.1	0.669017	0.419222	0.319607
88	0.2	0.4	0.1	0.1	0.2	0.667967	0.421553	0.32155
89	0.2	0.4	0.1	0.2	0.1	0.666153	0.425123	0.323693
90	0.2	0.4	0.2	0.1	0.1	0.668324	0.419591	0.320308
91	0.2	0.5	0.1	0.1	0.1	0.666658	0.421094	0.322317
92	0.3	0.1	0.1	0.1	0.4	0.670691	0.419576	0.318897
93	0.3	0.1	0.1	0.2	0.3	0.670381	0.423344	0.320045
94	0.3	0.1	0.1	0.3	0.2	0.670085	0.424922	0.320559
95	0.3	0.1	0.1	0.4	0.1	0.668623	0.426133	0.321637
96	0.3	0.1	0.2	0.1	0.3	0.672404	0.417947	0.316783
97	0.3	0.1	0.2	0.2	0.2	0.671988	0.421622	0.317911
98	0.3	0.1	0.2	0.3	0.1	0.670661	0.423278	0.31906
99	0.3	0.1	0.3	0.1	0.2	0.672474	0.417538	0.316524
100	0.3	0.1	0.3	0.2	0.1	0.67101	0.421252	0.318293
101	0.3	0.1	0.4	0.1	0.1	0.671181	0.417076	0.317189
102	0.3	0.2	0.1	0.1	0.3	0.668294	0.420226	0.320983
103	0.3	0.2	0.1	0.2	0.2	0.667509	0.42372	0.322523
104	0.3	0.2	0.1	0.3	0.1	0.666099	0.425242	0.323755
105	0.3	0.2	0.2	0.1	0.2	0.669639	0.418613	0.319188
106	0.3	0.2	0.2	0.2	0.1	0.668032	0.42211	0.321117
107	0.3	0.2	0.3	0.1	0.1	0.668715	0.417923	0.319459
108	0.3	0.3	0.1	0.1	0.2	0.667266	0.420332	0.321568
109	0.3	0.3	0.1	0.2	0.1	0.665416	0.42374	0.32396
110	0.3	0.3	0.2	0.1	0.1	0.667585	0.418445	0.320548
111	0.3	0.4	0.1	0.1	0.1	0.665762	0.419918	0.322587
112	0.4	0.1	0.1	0.1	0.3	0.669714	0.418928	0.319515
113	0.4	0.1	0.1	0.2	0.2	0.66908	0.422522	0.320928
114	0.4	0.1	0.1	0.3	0.1	0.667615	0.424075	0.322189
115	0.4	0.1	0.2	0.1	0.2	0.671176	0.417289	0.317571
116	0.4	0.1	0.2	0.2	0.1	0.669624	0.420874	0.319442
117	0.4	0.1	0.3	0.1	0.1	0.670178	0.416604	0.31787
118	0.4	0.2	0.1	0.1	0.2	0.667399	0.41951	0.321435
119	0.4	0.2	0.1	0.2	0.1	0.665551	0.422885	0.323629
120	0.4	0.2	0.2	0.1	0.1	0.667724	0.417649	0.32018
121	0.4	0.3	0.1	0.1	0.1	0.665498	0.419252	0.322568
122	0.5	0.1	0.1	0.1	0.2	0.668947	0.418478	0.319955
123	0.5	0.1	0.1	0.2	0.1	0.667151	0.421974	0.322088
124	0.5	0.1	0.2	0.1	0.1	0.669331	0.416589	0.318576
125	0.5	0.2	0.1	0.1	0.1	0.6658	0.41871	0.322172
126	0.6	0.1	0.1	0.1	0.1	0.667286	0.417811	0.320784

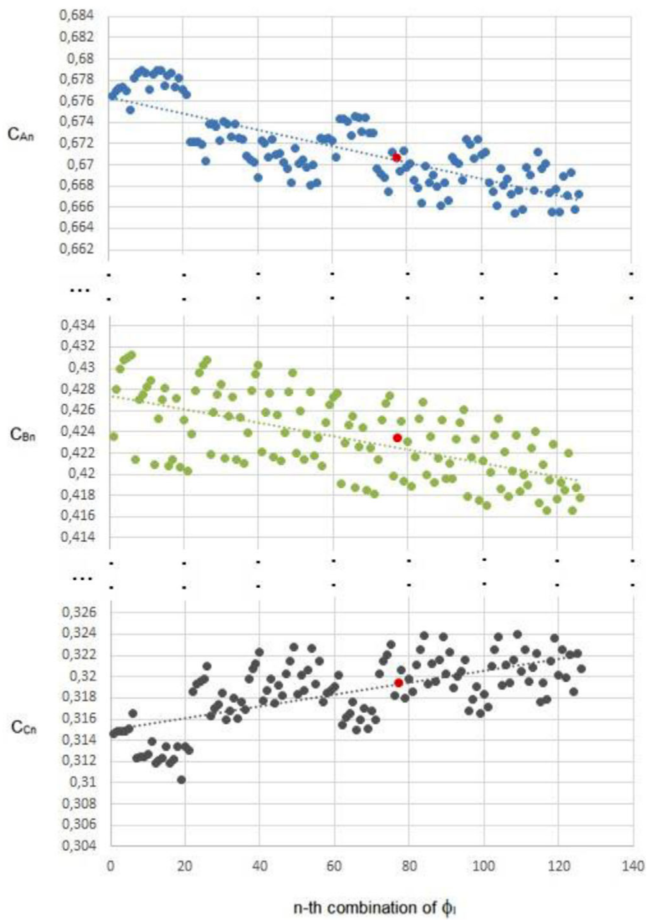

FIGURE A1 Dispersion of the proximity coefficients (C_{in}) as the 'Opinion Value' coefficient (ϕ_1) varies (paradoxical scenario)

TABLE A8 Descriptive analysis of the variable 'C_{in}' (paradoxical scenario)

Variable	Obs	Mean	SD	Min	Max
C_{An}	126	0.6715229	0.0035611	0.665416	0.678922
C_{Bn}	126	0.4234147	0.0037604	0.4165892	0.4312694
C_{Cn}	126	0.3185145	0.003144	0.3103	0.32396

The proximity coefficients, as the opinion value changes, are ordered starting from the left, by the configuration with the utmost importance of the opinion of the decision makers of the administrative-governance area ($\phi_1 = 0.1$, $\phi_2 = 0.1$, $\phi_3 = 0.1$, $\phi_4 = 0.1$, $\phi_5 = 0.6$) to the configuration, increasing to the right, with the utmost importance of the opinion of the clinical decision makers ($\phi_1 = 0.6$, $\phi_2 = 0.1$, $\phi_3 = 0.1$, $\phi_4 = 0.1$, and $\phi_5 = 0.1$).

Also in this case, the red points indicate the proximity coefficients values obtained considering the configuration in which the DMs have the same importance ($\phi_1 = 0.2$, $\phi_2 = 0.2$, $\phi_3 = 0.2$, $\phi_4 = 0.2$, and $\phi_5 = 0.2$). The results of the descriptive analysis of the variable C_{in} are shown in Table A8.

Although the standard deviation of the C_i of alternative A is no longer the lowest, it remains the most robust in terms of variability, as evidenced by the calculation of the coefficients of variation of each alternative's proximity coefficient (CV_i) shown below: $CV_A = 0.53029727$; $CV_B = 0.88811276$; and $CV_C = 0.98708222$. Again, the order is CV_A – CV_B – CV_C.