Received 22 September 2023
Revised 1 March 2024
Accepted 10 April 2024

Realising the promise of value-based purchasing: experimental evidence of medical device selection

Juri Matinheikki, Katie Kenny and Katri Kauppi
Department of Information and Service Management, Aalto University School of Business, Espoo, Finland
Erik van Raaij
Erasmus University Rotterdam, Rotterdam, The Netherlands, and
Alistair Brandon-Jones
School of Management, University of Bath, Bath, UK

Abstract

Purpose – Despite the unparalleled importance of value within healthcare, value-based models remain underutilised in the procurement of medical devices. Research is needed to understand what factors incentivise standard, low-priced device purchasing as opposed to value-adding devices with potentially higher overall health outcomes. Framed in agency theory, we examine the conditions under which different actors involved in purchasing decisions select premium-priced, value-adding medical devices over low-priced, standard medical devices.

Design/methodology/approach – We conducted 2 × 2 × 2 between-subjects scenario-based vignette experiments on three UK-based online samples of managers (n = 599), medical professionals (n = 279) and purchasing managers (n = 449) with subjects randomly assigned to three treatments: (1) cost-saving incentives, (2) risk-sharing contracts and (3) stronger (versus weaker) clinical evidence.

Findings – Our analysis demonstrates the harmful effects of intra-organisational cost-saving incentives on value-based purchasing (VBP) adoption; the positive impact of inter-organisational risk-sharing contracts, especially when medical professionals are involved in decision-making; and the challenge of leveraging clinical evidence to support value claims.

Research limitations/implications – Our results demonstrate the need to align incentives in a context with multiple intra- and inter-organisational agency relationships at play, as well as the difficulty of reducing information asymmetry when information is not easily interpretable to all decision-makers. Overall, the intra-organisational agency factors strongly influenced the choices for the inter-organisational agency relationship.

Originality/value – We contribute to VBP in healthcare by examining the role of intra- and inter-organisational agency relationships and incentives concerning VBP (non-) adoption. We also examine how the impact of such mechanisms differs between medical and purchasing (management) professionals.

Keywords Behavioural supply management, Agency theory, Purchasing, Healthcare, Experiments
Paper type Research paper

International Journal of Operations & Production Management
Vol. 44 No. 13, 2024
pp. 100-126
Emerald Publishing Limited
0144-3577
DOI 10.1108/IJOPM-09-2023-0779

The authors acknowledge the financial support of the Research Council of Finland (Grant Number 316020) for this research.

1. Introduction

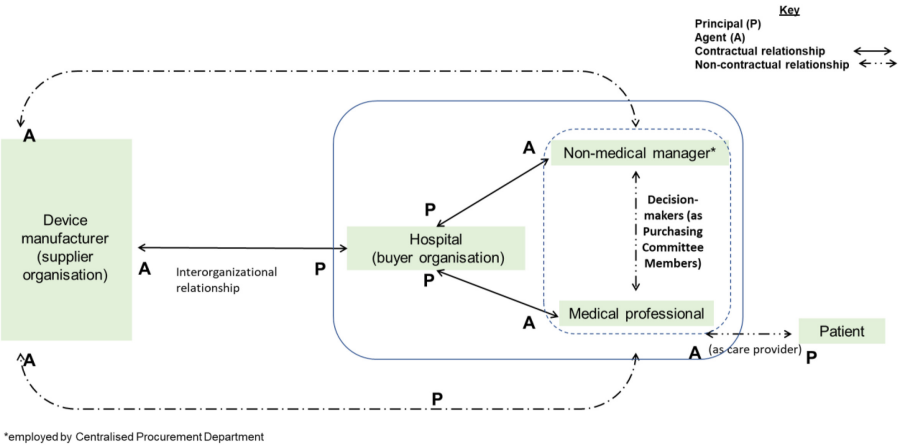
The value of purchasing is inherently vital to the healthcare industry (Grundy, 2016; Dobrzykowski, 2019; Lee *et al.*, 2020). Its critical nature was highlighted throughout the coronavirus disease 2019 (COVID-19) pandemic, where systemic breakdowns resulted in limited access to life-saving supplies and predatory practices regarding their procurement (Harland *et al.*, 2021). The predominant focus of value-based healthcare has been on payment models between the payer and the care provider (Prada, 2016). While the pandemic was an exceptional event and prioritising health outcomes in medical device procurement may seem obvious, less attention has been paid to medical device and technology procurement (Obremskey *et al.*, 2012; Pedroso *et al.*, 2022) despite their role in improving both the effectiveness for patients and efficiency of health systems (Robinson, 2008; European Commission, 2021).

The value-based purchasing (VBP) approach emphasises that higher quality, innovative devices used in diagnosis and care may improve health outcomes, in terms of clinical outcomes and improved patient experience, as well as offering long-term cost savings when compared to devices selected via price-based evaluations (Robinson, 2008; Sorenson and Kanavos, 2011; Prada, 2016). As a general principle, VBP requires buyers to consider elements of costs, and other sacrifices incurred, relative to benefits throughout a product lifecycle (Dumond, 1996; Meehan *et al.*, 2017; Kähkönen and Lintukangas, 2018; Gray *et al.*, 2020; Grundy, 2016), necessitating an expansion of the criteria typically used to select suppliers (Sorenson and Kanavos, 2011; Kokshagina and Keränen, 2022). In defining the "value" for value-based procurement, we refer to Gray *et al.*'s (2020) "total value contribution" approach to purchasing. We thus see value in our study context as the long-term value for the organisation from value-based devices, considering customer (in this case patient) value and looking at the life cycle contributions of value.

As procurement in the healthcare sector has become more professionalised and "economic" focused (Sanderson *et al.*, 2019), we need better adaptations of private sector strategies, such as the total value approach, to public health (Helper *et al.*, 2021). Yet extant literature suggests a strong preference for a price-based orientation in purchasing decision-making (Töytäri *et al.*, 2017; Kienzler, 2018). Purchasers are also responsible for balancing budgets in public health to keep pace with the high demand for care (Vogus *et al.*, 2020). Though a cost orientation has been heavily criticised in healthcare for several decades (Meehan *et al.*, 2017), alternative approaches are inconsistently applied even when outcome-based healthcare is a stated goal (Hurst *et al.*, 2019).

Medical device purchasing decisions are further complicated by the fact that they are often made by multi-disciplinary committees, where participants from different functions may have conflicting goals, partly due to professional backgrounds or formal and informal incentive structures (Atilla *et al.*, 2018; Hinrichs-Krapels *et al.*, 2022; Miller *et al.*, 2019). Those participating in medical purchasing committees function as, essentially, agents on behalf of the hospital organisation. This creates a classical *intra-organisational* agency problem, wherein the hospital needs to motivate the agents to act in their interest – in this case, the use of VBP. The context of VBP, however, is also related to an *inter-organisational* agency relationship, between the hospital as a buyer and the medical equipment supplier. We thus use agency theory (Eisenhardt, 1989; see Matinheikki *et al.*, 2022 for a supply chain specific review) to frame our study to better understand decision-making related to value-based procurement. Figure 1 captures the different multi-tiered intra- and inter-organisational agency relationships concerning the value-based procurement of medical devices and the decision-makers we study.

In this study, our goal is to understand how the decision-making priorities of the individuals who act as agents of the hospital within the context of the principal-agent relationship between the hospital and supplier may differ, and to what extent conditions in



Source(s): Created by authors

Figure 1.
Inter- and intra-organisational agency relationships in medical device purchasing

the intra- and inter-organisational agency relationship impact these priorities. Specifically, we aim to understand: (1) *Under which conditions decision-makers choose a premium-priced value-adding medical device over a low cost, standard medical device* and (2) *How factors related to the intra-organisational agency relationship of the decision-maker and factors related to the inter-organisational relationship between the buyer and supplier impact decisions.* We specifically address three factors: (1) cost-saving incentives, (2) risk-sharing contracts based on outcomes and (3) clinical value evidence.

Our research leverages a scenario-based vignette experiment (Rungtusanatham *et al.*, 2011) asking participants to make a procurement decision regarding cardiac implantable electronic devices (CIEDs), commonly known as pacemakers. We collected three different samples of general managers, medical professionals and purchasing managers – key stakeholders typically participating in purchasing committees. Whilst it is important to note that purchasing decisions typically occur through *interactive* committee work, our research explores *individual* decision-making to better understand conflicting goals and different types of risks between the multiple intra-organisational (organisation – employee) and the inter-organisational (buyer-supplier) relationships. As such, individuals serve as the basis for multilevel theorisation, as their decisions combine to form part of the overall sourcing decision (Carter *et al.*, 2015; Meschnig *et al.*, 2018). In the same way, individual performance has an impact on the overall performance of the buyer-supplier relationship, yet research focusing on these links is scarce (Oosterhuis *et al.*, 2005). As such, addressing this issue is a key contribution to our work.

Our study provides contributions to the nascent value-based domain in purchasing and supply chain management (Gray *et al.*, 2020) and more particularly to VBP in healthcare (Robinson, 2008; Lee *et al.*, 2020) by providing an increased understanding of the role of personal and organisational incentives and information asymmetry to purchasing decision-making as framed through agency theory. While agency theory is typically utilised to understand *either* a single employee (agent) acting on behalf of an organisation (principal) *or* a relationship between a supplier organisation (agent) and buyer organisation (principal), we contribute by examining *both* levels of agency relationships as they impact value-based procurement decision-making at the individual level. There are multiple and complex agency relationships in healthcare with partly compatible and partly diverging goals and payment structures, which must be considered in motivating behaviour and change towards value-

based approaches (Conrad, 2015; Kokshagina and Keränen, 2022; Peltokorpi *et al.*, 2020). Thus, our study also strongly addresses the calls for multilevel theorisation into supply chain-related phenomena (Carter *et al.*, 2015; Touboulic and Walker, 2015).

2. Agency theory in the value-based procurement context - theoretical background and hypotheses

As suggested by Kaczmarek (2017), the VBP context can be described as a multi-tiered agency configuration – a complex set of intra- and inter-organisational agency relationships taking place at the boundaries of functions and organisations. Our VBP context is tied to both an intra-organisational agency relationship between the hospital (principal) and the (purchasing) managers and medical professionals (agents) in the committee making the purchasing decisions, and to an inter-organisational agency relationship between the hospital (principal) and the supplier (agent). Both agency relationships can suffer from goal incongruence and information asymmetry, and have different aspects of agency risk, which can impact the behaviour of the decision-maker.

Given the two levels of agency relationships, there are two potential levels for the classical agency problem to arise when a task is delegated under conditions of goal incongruence and information asymmetry (Eisenhardt, 1989). However, agency theory is concerned with not only the commonly referenced "agency problem" but also with the less commonly discussed "problem of risk sharing," which *"arises when the principal and agent have different attitudes toward risk"* and may as a result prefer different actions (Eisenhardt, 1989, p. 58).

In our context of VBP, risks are also present in the different agency relationships (as per Figure 1), with the hospital facing short- or long-term financial risks depending on the product choice, the decision-makers facing monetary risks related to their compensation schemes, and medical professionals facing health risks for their patients [1]. The risk-sharing problem has an impact on the agency problem (at both levels) and on the actions of individual decision-makers. Figure 2 presents the agency problems and risk-sharing problems within the multi-tiered agency relationships that operate in this context.

We draw from agency theory, but our assumption of agent behaviour rests more on self-interest, rather than opportunism, i.e. self-interest *with guile* (which is indeed the term both Eisenhardt, 1989; Matinheikki *et al.*, 2022 also use in discussing the assumptions of agency

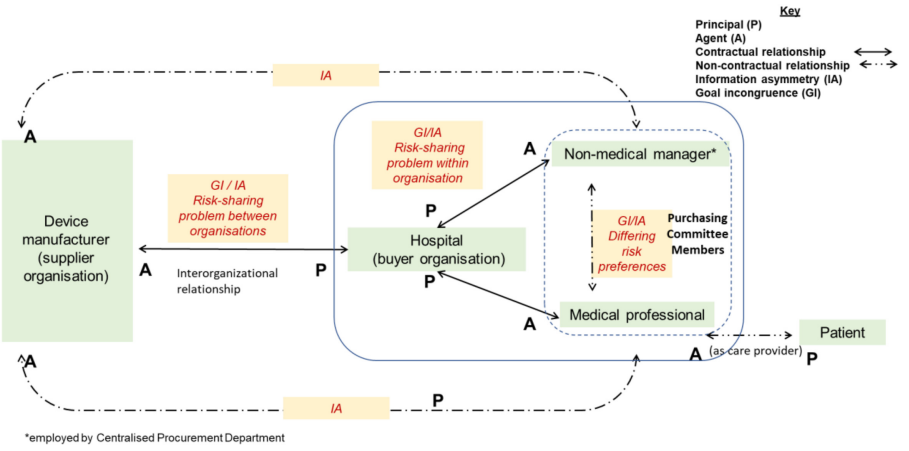


Figure 2.
The agency problems and risk-sharing problems within the intra- and inter-organisational agency relationships related to VBP

Source(s): Created by authors

theory). In contrast, the alternative behavioural assumption of *stewardship* (Davis *et al.*, 1997) is typical in the healthcare and public administration context (Kauppi and van Raaij, 2015). Medical professionals can be considered stewards because of their embedded relationships within hospitals and direct relationships with patients. Yet, this stewardship role has, to some degree, been altered through policy pressures to reduce costs, even in specialty areas such as cardiology (Miller *et al.*, 2019).

2.1 Decision-making incentives within the intra-organisational agency relationship between hospital and purchasing decision-makers

All the stakeholders participating in a purchasing committee act as agents on behalf of the hospital. This creates a classical agency problem, wherein the principal needs to motivate the agent to act in their interest through, for example, compensation systems (Jensen and Heckling, 1995). A large agency-based body of research demonstrates that incentive pay is a powerful mechanism to align agent interests with organisational goals (Tosi *et al.*, 1997; Gottschalg and Zollo, 2007; Villena *et al.*, 2018). Functional performance measures typically fall under either profit or cost-based rules, with the success of procurement departments and personal bonuses being predominantly based on savings targets (Hofmann *et al.*, 2014, p. 23). Supply chain managers have been shown to avoid high-risk decisions if they entail a personal compensation risk, regardless of the potential overall organisational benefits associated with such decisions (Villena *et al.*, 2018).

Similarly, in a healthcare context, misaligned incentives between key players in purchasing decisions have been noted (McKone-Sweet *et al.*, 2005; Lonsdale and Watson, 2005). Hospital administration and procurement functions have come to value cost savings (Helper *et al.*, 2021 or Sanderson *et al.*, 2019), while medical professionals are ultimately responsible for clinical outcomes and patient safety (Atilla *et al.*, 2018). This conflicts with medical professionals' incentives that are contingent on organisational structure. For instance, Abdulsalam *et al.* (2018) found that hospitals with higher levels of salaried physicians spend less on medical devices and supplies compared to hospitals with high levels of independent practitioners. The authors argue that this is due to better managerial control over employed physicians which curbs the agency's problem of excessive device spending. As such, whilst clinical outcomes matter to physicians, they are still subject to increasing cost savings pressures (Abdulsalam and Schneller, 2019; Vogus *et al.*, 2020).

Outcome-based contracts are often studied as a way to align incentives with healthcare provider goals (Sanderson *et al.*, 2019). When making decisions, purchasers are often paired with medical professionals who are both direct providers of healthcare as employed by the hospital (principal) and knowledgeable consultants of which devices they would like purchased (agents). Medical professionals play a dual role within this relationship system, where they are both incentivised to reduce costs and deliver quality care to patients. Vogus *et al.* (2020) discuss the pressures medical professionals are under to deliver the best care at the lowest cost.

As such, it seems likely that medical professionals would also be responsive to cost-saving incentives tied to supply savings but potentially to a lesser extent than general or purchasing managers as the medical professionals are more influenced by the health risks due to their agency relationship with the patient. These aspects lead us to hypothesise the following:

H1a. Organisational cost-saving incentives decrease managers' likelihood of choosing premium-priced, value-adding products.

H1b. Organisational cost-saving incentives decrease medical professionals' likelihood of choosing premium-priced, value-adding products.

H1c. Organisational cost-saving incentives decrease purchasers' likelihood of choosing premium-priced, value-adding products.

2.2 Inter-organisational agency relationship between hospital and supplier

We also address the agency relationship *between* organisations in the VBP context, between a buyer and a supplier. This relationship also suffers from a potential agency problem of goal incongruence and information asymmetry and is impacted by the risk-sharing problem of the parties. We will discuss two potential mechanisms that the decision-maker can use to mitigate the inter-organisational agency problem: (1) a risk-sharing contract, which tackles both the goal incongruence as well as the risk-sharing problem and (2) medical value evidence, which is in essence a signalling mechanism employed by the agent to counter adverse selection and reduce information asymmetry regarding product performance. However, as noted earlier concerning the multilayered agency context, decisions regarding the inter-organisational contract are made by decision-makers who are also subject to the agency problem and risk-sharing problem within their own intra-organisational context.

2.2.1 Risk-sharing contract. In purchasing, goal incongruence between the buyer and the supplier is traditionally solved through outcome-based contracts (Matinheikki *et al.*, 2022). Risk-sharing mechanisms are designed to share the burden of loss as well as reduce moral hazard by ensuring the supplier bears partial responsibility for quality-related incidents (Zu and Kaynak, 2012; Tse *et al.*, 2018). In the context of healthcare, the general approach is to include a certain element of risk-sharing related to a targeted health outcome in the medical device purchasing contracts. Thus, when discussing risk-sharing contracts, we refer to how the consequences are financially shared rather than the detailed obligations and duties to mitigate those risks.

Such performance-based risk-sharing arrangements (from now on referred to as risk-sharing contracts) are increasingly used in medical device purchasing (Antonanzas *et al.*, 2019). Offering a risk-sharing contract is an increasingly leveraged strategy in so-called value-based selling since it can reduce the buyer's ambiguity aversion (Töytäri *et al.*, 2017). Similar contracts have proved optimal in healthcare provision, in which payments are adjusted if adverse patient outcomes emerge (Fuloria and Zenios, 2001). As such, this can be seen as having a dual guarantee, creating an effective selling point: the supplier's belief in the reduced number of medical complications as well as the promise to split any costs if failures do occur (in essence mitigating moral hazard and adverse selection agency problem as well as the risk-sharing problem simultaneously). The challenge in such risk-sharing contracts is typically related to the measurement of highly uncertain outcomes as well as defining how much risk can be allocated to the device manufacturer. In well-established and relatively simple medical interventions, this is possible but when treatment requires input from multiple parties (e.g. a complex surgical implantation), allocating the risks and responsibilities becomes difficult.

The above discussion characterises the role of the risk-sharing contract in reducing the risk for the hospital in the inter-organisational agency relationship. However, from an individual decision-making perspective, it is also relevant to discuss how different stakeholders value risk-sharing contracts especially if they are purely tied to clinical outcomes.

Medical professionals often have a low tolerance for risk concerning the health outcomes of their patients (Cook *et al.*, 2008). Therefore, any signals about high care quality (such as willingness to share risks) are often perceived positively among medical practitioners (Cook *et al.*, 2008). Purchasing managers also tend to be risk averse (Kirchoff *et al.*, 2016), and concerned as to whether the superior value in exchange for a higher price will materialise

(Anderson and Wynstra, 2010), and given a public health setting, are likely more sensitive to health risks. Thus, the purchasing managers will be more likely to enter into a more expensive contract when the risk of adverse selection is reduced through a risk-sharing mechanism. The benefits of a higher-value alternative, however, are often felt by a different organisational function, with purchasing managers bearing the consequences of the higher price in their personal incentives without receiving credit for improved outcomes further down the line (Anderson and Wynstra, 2010; Anderson *et al.*, 2000; Lonsdale and Watson, 2005), nor in this case for the potential risk-sharing payments. These aspects taken together; we hypothesise the following:

H2a. A risk-sharing contract offered by the supplier increases managers' likelihood of choosing premium-priced, value-adding products.

H2b. A risk-sharing contract offered by the supplier increases medical professionals' likelihood of choosing premium-priced, value-adding products.

H2c. A risk-sharing contract offered by the supplier increases purchasers' likelihood of choosing premium-priced, value-adding products.

2.2.2 Medical value evidence. In medical device purchasing, information asymmetry exists between the manufacturer and all decision-makers within the hospital. This forces the manufacturer to signal the value created by their devices, i.e. the agent needs to signal their capabilities to the principal (Bergen *et al.*, 1992). Clinical evidence is a way to reduce such information asymmetry and signal healthcare product value. Device manufacturers are encouraged to provide compelling medical evidence to support the value proposition of value-adding devices (Hurst *et al.*, 2019). For instance, in Grundy's (2016) case study of VBP, physician experience appeared to be the primary evidence used by managers when selecting new devices, even if these devices are more expensive. Clinical evidence can thus reduce the risk of adverse selection for the hospital.

As decision-makers tend to be risk averse, the type of product information presented can also heavily influence their choices (Wong, 2023). Several studies indicate that purchasing managers are inclined to choose higher-value options given credible evidence (Riedl *et al.*, 2013), relying on formal market research and expert recommendations to reduce decision-making uncertainty (Formentini and Romano, 2016; Chae *et al.*, 2019).

Evidence-based outcomes are also pivotal for value-based healthcare decisions (Bae, 2015; Hurst *et al.*, 2019). However, despite medical evidence upskilling and value focus, evidence may not drive decision-making as strongly as one might anticipate (Hurst *et al.*, 2019), as more information does not necessarily reduce uncertainty if receivers do not know what to do with it (Hurst *et al.*, 2019). Lacking the ability to interpret health information increases decision-maker uncertainty and ambiguity (Gardner *et al.*, 2015; Miller *et al.*, 2019) and hinders the ability for decision-makers, especially those without medical education, to assess the total expected value of alternatives (Obremskey *et al.*, 2012; Riedl *et al.*, 2013). In the VBP context, information asymmetry can thus exist within the hospital and purchasing committee; medical professionals are often better informed about clinical performance of devices (Atilla *et al.*, 2018) than, for example, purchasing managers or hospital administration. During the past decades, significant effort has been put into training medical professionals' competence to review such evidence, which is often labelled as a core competence in evidence-based medical practice (Albarqouni *et al.*, 2018). This is why medical professionals are valued members of purchasing committees but also why they often have a relative power advantage in driving decisions; and thus, an opportunity to, for example, choose devices based on their own preference rather than medico-economic criteria. We hypothesise the following:

H3a. The availability of stronger clinical evidence of the benefits of the value-adding product increases the managers' likelihood of choosing premium-priced, value-adding products.

H3b. The availability of stronger clinical evidence of the benefits of the value-adding product increases the medical professionals' likelihood of choosing premium-priced, value-adding products.

H3c. The availability of stronger clinical evidence of the benefits of the value-adding product increases the purchasers' likelihood of choosing premium-priced, value-adding products.

3. Experimental research design

We conducted three scenario-based vignette experiments (Rungtusanatham *et al.*, 2011; Eckerd *et al.*, 2021) via Prolific to collect UK-based samples. Prolific is recognised for better data quality when compared to other online data platforms such as Mturk (Peer *et al.*, 2022). To test our three hypotheses, we adopted a $2 \times 2 \times 2$ between-subjects factorial design in which subjects were randomly assigned to three treatments: (1) incentive, (2) risk sharing and (3) evidence, resulting in eight experimental conditions.

The vignette was inspired by Medtronic's TYRX cardiac envelope for pacemakers [2]. Medtronic built a research program, which included a major scale global randomised controlled trial (RCT) indicating the clear drop in the infection rates of TYRX patients (Tarakji *et al.*, 2019). In the initial stages of TYRX market entry, Medtronic offered risk-sharing contracts, including paybacks to hospitals for each post-operation infections, to boost the widespread deployment of TYRX envelopes.

Participants were tasked with a pacemaker purchasing decision between renewing a contract for either a standard pacemaker or a value-adding enveloped pacemaker. These two choices varied in terms of unit price, post-operation infection risks, infection-based medical costs and death rate. All values followed either the industry reports (Wenzl and Mossialos, 2018) or scientific studies (Ahmed *et al.*, 2019; Tarakji *et al.*, 2019) in the UK. For the general manager and purchasing manager samples, participants were given the role of a procurement specialist for a public health purchasing agency in the UK. In the medical professional sample, the participant was given the role of being a member of a medical purchasing committee (see Appendix A for vignette description).

Recognising the potential limitations of online experiments, we took care to (1) pre-screen participants, (2) offer generous compensation contingent on passing three attention checks (Abbey and Mellony, 2017) and (3) make choices financially consequential using bonus payments. Additionally, we pre-tested the experimental design with a generic sample to ensure research quality. We then collected three different samples of the types of professionals who participate in medical device purchasing decisions: general managers, medical professionals and purchasing managers.

3.1 Experimental manipulations

Under the *incentive condition*, participants were presented an organisational scheme in which they would gain a monetary bonus when they achieved cost savings compared to the previous supply contract (cost differences for standard and value-adding devices were £2,200 vs £3,000 per unit, giving an £800,000 difference for quoted 1,000 units). If choosing a standard contract, participants would gain a £1,000 in-scenario bonus, which corresponded to an extra £1 in their participation fee.

Under the *risk-sharing condition*, participants were told the manufacturer would share 50% of the medical costs in case of infection if they chose the enveloped device. These costs

were reported to be on average £140,000 for 1,000 patients with the reported infection rates yielding potential cost savings of £70,000. This corresponds to the actual risk-sharing strategy used by Medtronic [3].

Finally, the *evidence condition* manipulated the source strength of medical evidence on which the infection rates were based. In the treatment groups of all samples, the evidence was provided by an independent medical research institute study based on a global RCT. Such trials are typically accepted as the strongest or the second strongest level of medical evidence [4] and as such reduce the uncertainty of the enveloped pacemaker's value. In control groups of samples 1 and 2, the evidence was claimed to be based on a clinical trial in one hospital but as explained later, we witnessed poor levels of manipulation. Therefore, in the third sample, we dropped any notion regarding the strength of evidence in the control group and relied only on the source. In all samples and conditions, it was claimed that the information was brokered by the device manufacturer.

After the subject had decided, each manipulation was followed by two specific checks to avoid demand effects (Lonati *et al.*, 2018). The incentive condition had two extra checks, which were placed before and within the vignette to ensure that the participants realised that their decisions would result in real payoffs and that they had understood the incentive structure. Evidence manipulations were also slightly different since both treatment and control conditions provided some evidence with varying degrees of trustworthiness. The order of choices was randomised in each check.

3.2 Control variables

Since our scenario involved a somewhat high cognitive load, it can be assumed that subjects' cognitive capability and especially their tendency to reflect before making decisions may affect their final decision. Therefore, we used the cognitive reflection test (CRT) to control for subject's reflective thinking (Frederick, 2005; Toplak *et al.*, 2011). CRT consists of three questions each having an intuitive but incorrect answer and the correct answer requires reflection. In Operations and Supply Chain Management (OSCM), high results in CRT have been found to highly correlate with, for example, decision-making regarding medical device recalls (Ball *et al.*, 2018). To avoid inflation, we used one item from the original test by Fredrick (2005) and two questions from a more recent expansion by Toplak *et al.* (2011). We transformed the three CRT questions into a binary "high CRT" variable, which takes value one if two out of three questions were correct (cf. Ball *et al.*, 2018).

Risk perception can be a strong determinant of decision behaviour (Kull *et al.*, 2014). Situational risk perception is often affected by risk attitude, ranging from risk aversion to risk seeking. Thus, we used the Domain Specific Risk Taking (DOSPERT) scale's three-item financial investment risk scale to control for subjects' risk-taking attitudes in investment decisions (Blais, 2006). In addition to risk attitude, we examined the post-decision perception of the riskiness of both device choices on a 1 (Not risky at all) to 7 (Extremely risky) response format, which is a modification of a single-question risk perception scale (Ganzach *et al.*, 2008). However, we used it only for post-hoc checks and not as a control variable in the regression models since it could be affected by our manipulations. Furthermore, long-term orientation (LTO) has also been found to significantly affect managerial decision-making and risk preferences, thus, we used Bearden *et al.* (2006) four-item scale on time orientation in planning.

Lastly, we collected detailed demographic information to increase statistical precision and gain information on our sample and the effectiveness of the randomisation. These included age, management experience, healthcare experience, procurement experience, gender, size of current organisation, industry, employment sector, current work role in procurement and time used to conduct the study. These controls were in line with recent purchasing-oriented

behavioural operations research (Ball *et al.*, 2018; Yan *et al.*, 2018). Additionally, we inquired if the subjects or any of their close relatives or friends had a pacemaker. All control variables were collected after the experimental scenario to avoid giving out false signals that, e.g. we would value long-term or less risky choices.

3.3 Description of analysis

Our analysis strategy builds on calculating average treatment effects (ATEs) for all eight experimental conditions and then calculating the main effects for the three hypothesised conditions of incentive, risk-sharing and evidence (i.e. the average difference across the three sub-conditions in which the main category is present). ATE in our case is the difference in the proportion of value-adding device choices when compared to the no-manipulation baseline. However, we wanted to add covariates, which may help in improving precision by controlling for potential condition imbalance (Wang *et al.*, 2017). Therefore, our main analysis uses a hierarchical Ordinary Least Squares (OLS) regression analysis to evaluate our experimental effect without (Model 1) and with (Model 2) control variables.

Despite having a binary outcome variable (device choice), we chose a linear OLS model over other alternatives (such as logistic regression) since it allows robust estimation (when using heteroskedastic standard errors) and a clearer interpretation of the results (Gomila, 2021). In this case, regression coefficients of the three hypothesised main effects are comparable to percentage point differences in the proportion of value-adding device choices. As a robustness check, we also ran the same analysis using logistic regression with heteroskedastic standard errors and observed no major differences in the results. These analyses are available in Appendix D.

In addition, to correctly estimate the three main effects and account for the interaction effects, we coded our three experimental treatments (incentive, risk sharing and evidence) as contrasts (Wiens and Nilsson, 2017). For ease of interpretation, we decided to use contrast weights of −0.5 and 0.5, which practically means that instead of coding experimental variables as 0/1 (i.e. a "dummy coding"), we coded them as −0.5/0.5 i.e. an "effects coding" (Bech and Gyrd-Hansen, 2005). This ensures that in Model 1 and Model 2, the regression coefficients of the main effects for incentive, risk sharing and evidence variables are directly comparable to the main effects calculated via the traditional ATE estimation route. All the statistical analyses were conducted with Python and its multiple data analysis packages. The Python code is available from the authors upon request.

4. Sample 1: general managers

In the first sample, general managers were chosen to study how those not in the medical field would approach this purchasing situation (i.e. representing a more arm's length perspective). When recruiting the sample of general managers from Prolific, we adopted three pre-screening criteria: (1) the participant should reside in the UK, (2) be employed full-time (3) and have management experience. At the time of the study, Prolific had a sample pool of about 7,800 subjects meeting these criteria. Based on the power analysis of our pilot study (see Appendix B), to ensure sufficient statistical power for the strongest effect (incentive = −11.5, $n_{\text{treat}} = 30$, $n_{\text{total}} = 62$), we needed to attain a sample size of 572 subjects. To mitigate possible risks in data collection, we decided to split the data collection in two rounds (controlled for in our subsequent analyses).

We obtained usable data for a total of 599 practising managers out of 731 subjects who started the experiment (i.e. those who passed all three attention checks, fully completed the task, and were eligible for the payment of the fixed fee [£4]). This sample size allowed

having on average 75 subjects per experimental condition, well exceeding the typical guideline of "$n > 50$ per cell" (Lonati *et al.*, 2018, p. 20). The sample size eventually gave us 65% statistical power for the incentive effect, which ended up being slightly smaller than the original power analysis suggested (see Appendix C Table C1 for descriptive statistics and correlations).

4.1 Manipulation checks

First, under the incentive condition, most subjects understood that their decision may be consequential, 93% of subjects understood the in-scenario bonus scheme within the scenario and 81% recalled the conversion rate correctly post-hoc.

The risk-sharing manipulation appeared to work well. The same cannot be said for the evidence manipulation. In the first round of data collection, we observed a low pass rate in a single evidence check. We thought that this would be because subjects find it difficult to identify the correct answer with two similar options (clinical trial vs global RCT). Thus, in the second round, we split the check into two questions with condition-dependent choices in the latter question (Evd-check 2). As indicated in Table C2 (see Appendix C), these corrective measures did not help the subjects to distinguish between different sources of evidence after making the decision. In fact, the rate of correct answers dropped by nearly 20%-points in both groups. However, interestingly the low pass rates appeared to be stable between two manipulations (a two-tailed *t*-test between two groups yields insignificant results in all checks). We conclude that this indicates that our general manager sample may not be receptive to information of any kind regarding the medical evidence which may require specialised knowledge to differentiate between the sources of evidence. We will return to this argument in our other samples and in the discussion section.

Despite the low rate of evidence manipulation, we decided to keep subjects who did not pass manipulation checks. The practice of deleting subjects who failed manipulation checks has been shown to cause more damage than good since it will imbalance the experimental conditions inducing unpredicted biases to effect sizes (Aronow *et al.*, 2019). Therefore, we interpret the evidence manipulation effect conservatively. However, we ran the same analysis for a sub-sample of those who passed all checks as a robustness check. The results appear the same, though not statistically significant.

4.2 Results

Table 1 summarises the regression models for the general manager sample. The two leftmost columns show the results as unstandardised beta coefficients of Model 1a (the experimental effects and their interactions only) with standard errors and 95% confidence intervals. Only the incentive effect (H1a) appears significant at $p = 0.02$. The effect size indicates a high 9.1%-point decrease in the choice of value-adding premium-priced devices. Thus, our analysis supports H1a.

Adding the controls in Model 2a leads to a clear increase in R-squared ($=0.10$) but minimal changes in the estimates of independent variables, which indicates that the experimental conditions were balanced in terms of control variables. Interestingly, high CRT scores appear to be strongly associated with choosing the standard device (effect size of -15%-points). Similarly, but to a lesser extent, the subjects scoring high on financial risk-taking (DOSPERT) are less likely to choose the value-adding option. Respondents who know someone with a pacemaker have an increased likelihood of value-based choices by eleven percentage points. Other control variables were not statistically significant.

5. Sample 2: medical professionals

We collected the second sample targeting medical professionals, who we anticipated would have better subject matter expertise to evaluate the strengths of different types of medical

Variables	Model 1a β coefficients (std. errors)	95% CI	Model 2a β coefficients (std. errors)	95% CI
Intercept	0.67*** (0.02)	[0.63, 0.71]	0.94***	[0.53, 1.35]
Incentive	−0.091** (0.04)	[−0.17, −0.02]	−0.083** (0.04)	[−0.16, −0.01]
Risk sharing	0.064 (0.04)	[−0.01, 0.14]	0.057 (0.04)	[−0.02, 0.13]
Evidence	0.036 (0.04)	[−0.04, 0.11]	0.032 (0.04)	[−0.04, 0.11]
Incentive* Risk sharing	−0.099 (0.08)	[−0.25, 0.05]	−0.085 (0.08)	[−0.24, 0.07]
Incentive* Evidence	−0.026 (0.08)	[−0.18, 0.13]	−0.016 (0.08)	[−0.17, 0.14]
Risk sharing* Evidence	0.027 (0.08)	[−0.12, 0.18]	0.003 (0.08)	[−0.15, 0.16]
Incentive* Risk sharing* Evidence	0.133 (0.16)	[−0.17, 0.44]	0.138 (0.16)	[−0.17, 0.45]
High CRT			−0.174*** (0.05)	[−0.26, −0.09]
Risk-taking attitude (DOSPERT)			−0.041*** (0.01)	[−0.07, −0.01]
Next of kin has a pacemaker			0.11** (0.04)	[0.01, 0.19]
Other controls[1]			included not significant (n.s.)	
Sample size	599		599	
R-squared	0.019		0.104	
F-statistic	1.788*		5.017***	

Note(s): *p-value <0.10.; **p-value <0.05.; ***p-value <0.01
[1](organisation size; industry; data collection round; procurement experience; age; gender; procurement position; procurement experience; management experience; LTO; time)
Source(s): Created by authors

Table 1.
General manager sample: OLS regressions of choice with heteroskedastic standard errors

evidence. We recruited an online sample of medical professionals from Prolific by using three pre-screening criteria: (1) a participant should reside in the UK, (2) be employed full-time, (3) work in medical/healthcare industry. Following the same procedure as our previous sample, we collected a final sample of 276 medical professionals (see Appendix C Table C3).

Approximately 30.1% of the subjects identified themselves as nurses, 22.8% as working in an administration role, 10.1% as allied health professionals [5], 8% as medical doctors, 5.4% as any type of clinical or medical scientists or researchers, 5.4% as clinical support specialists and 4.7% as a pharmacist. The remaining 13.4% were categorised as "other," involving various specialties such as psychologists, social workers, medical engineers, dentists, etc. Approximately 80.5% of subjects worked for a public employer, 17.7% for a private employer and 1.8% for an NGO. Additionally, 32% of participants claimed to work currently in a procurement-related position.

5.1 Manipulation checks

We used the same manipulation checks as the latter rounds of our general managers' sample. A similar pattern occurs whereby incentive and risk-sharing checks were passed relatively well, but we observe a low pass rate for the evidence manipulation check (see Appendix C Table C4). As such, our *a priori* assumption that medical professionals would be more responsive to evidence manipulation is not supported.

5.2 Results

The results of two OLS regression models are listed in Table 2. Model 1b estimates the main effects only and Model 2b adds the controls. The incentive effect remains insignificant in both models meaning that we do not find support for H1b. Interestingly, adding controls only slightly dilutes the risk-sharing and evidence effects of which both remain significant (at levels $p = 0.007$ and $p = 0.014$ respectively in Model 2b).

In evidence check 2 (asking the subject to identify the source of information), we witness a significantly higher pass rate in treatment (43%) than in control (28%) at level $p < 0.01$ (two-tailed *t*-test). The difference is even greater for the medical doctor sub-sample, who we assumed to have the highest literacy to interpret medical evidence (64% in treatment vs 13% in control, $p < 0.05$), but this can be explained due to the small sub-sample size (14 doctors in evidence treatment vs 8 doctors in control). This indicates that only the highest level of medical evidence (i.e. an RCT) can invoke the subject's attention to the source of evidence. We decided to continue our analysis without dropping the subjects but ran the post-hoc robustness check with the sample passing all the manipulation checks. Interestingly, the statistically significant effects persist, and the estimated effects grow larger.

Thus, we found support in this group for H2b indicating that a risk-sharing contract increases the tendency to opt for value-adding devices. The average increase in value-adding device choices is as high as 16%-points. We are more conservative regarding Hb given the apparent problems in evidence manipulation, albeit the effect appears strong (15%-points increase in Model 2). This time, we also witness a weakly significant interaction effect on incentive × risk sharing × evidence, which indicates that the presence of all conditions negatively moderates, i.e. dilutes the effect of experimental conditions.

Variables	Model 1b β coefficients (std. errors)	95% CI	Model 2b β coefficients (std. errors)	95% CI
Intercept	0.697*** (0.03)	[0.64, 0.75]	0.462* (0.26)	[−0.05, 0.98]
Incentive	−0.074 (0.06)	[−0.19, 0.04]	−0.078 (0.06)	[−0.2, 0.04]
Risk sharing	0.168*** (0.06)	[0.06, 0.28]	0.16*** (0.06)	[0.04, 0.28]
Evidence	0.133** (0.06)	[0.02, 0.24]	0.15** (0.06)	[0.03, 0.27]
Incentive* Risk sharing	0.028 (0.11)	[−0.19, 0.25]	0.025 (0.12)	[−0.21, 0.26]
Incentive* Evidence	0.089 (0.11)	[−0.13, 0.31]	0.144 (0.12)	[−0.09, 0.37]
Risk sharing* Evidence	0.102 (0.11)	[−0.12, 0.32]	0.026 (0.12)	[−0.2, 0.25]
Incentive* Risk sharing* Evidence	−0.373* (0.23)	[−0.82, 0.07]	−0.44* (0.25)	[−0.92, 0.04]
High CRT			−0.17*** (0.07)	[−0.3, −0.04]
Healthcare experience			−0.009* (0.01)	[−0.02, 0]
Other controls[1]			included not significant (n.s.)	
Sample size	276		276	
R-squared	0.073		0.197	
F-statistic	1.339		2.885***	

Note(s): *p-value <0.10.; **p-value <0.05.; ***p-value <0.01

[1](DOSPERT, next of kin has pacemaker, organisation size; industry; data collection round; procurement experience; age; gender; procurement position; LTO; medical role, sector, time)

Source(s): Created by authors

Table 2. *Medical professional sample:* OLS regressions of choice with heteroskedastic standard errors

Regarding the control variables, we witness a similar pattern to the manager sample with high CRT being associated with a lower proportion of value-adding device choices (−17%-points). Interestingly, experience in the healthcare sector is weakly significant (at $p = 0.09$) so one additional year of experience is associated with a 0.9%-point drop in value-adding device choices.

6. Sample 3: purchasing managers

Finally, we recruited a sample of subjects with management experience and specific responsibilities in purchasing and supply chain management using Prolific. We adopted four pre-screening criteria: (1) the participant should reside in the UK, (2) be employed full-time, (3) have management experience and (4) have supply chain/logistics-related decision-making responsibilities at work. Following the same procedure, we obtained a total of 449 responses who passed all three attention checks and were eligible for a fixed fee (£4). The subjects had on average 11 years (SD 8.2 years) of management experience, 87% worked in a procurement-related position and 11% in health care and social assistance during the time of the study (see Appendix C Table C5 for descriptive statistics and correlations).

6.1 Manipulation checks

Given the low pass rate of the evidence manipulation checks in the general managers' sample, we slightly altered both the manipulation itself and the corresponding checks in the purchasing manager sample. Instead of using two different evidence sources, we used no evidence provided vs evidence stemming from a global RCT. The purpose was to induce a stronger manipulation effect and allow subjects under treatment to identify the correct source when asked in the post-decision manipulation check. We dropped one comprehension check for the incentive manipulation to reduce potential priming effects and cognitive burden of the subjects.

Regarding incentive and risk manipulations, we observe slightly higher pass rates than for the general managers' sample (on average 87% and 96%, respectively) and can conclude that these manipulations work satisfactorily. However, we continued to observe a low, albeit slightly improved, pass rate (on average 60%) for the evidence manipulation. Even 24% of those now under control, who were provided no medical evidence, answered that they were provided medical evidence (see Appendix C Table C6).

6.2 Results

The results of two OLS regression models are listed in Table 3. In the manipulations only Model 1c, we observe a significant incentive effect at $p = 0.04$ and the effect size indicates a 9-percentage point decrease in the choice of value-adding premium-priced device. This again supports H1c. Interestingly, the risk-sharing contract has a negative point estimate in this sample, but this effect is not statistically significant. The evidence effect is weakly significant ($p = 0.09$) but given the problems in manipulation, this needs to be interpreted with caution.

In this sample, adding controls in Model 2 again leads to changes in the point estimates of our variables of interest. Most importantly, the incentive effect is diluted by 1.3%-points resulting in only weak statistical significance ($p = 0.08$) and partial support for H1c. This model does not support the other two hypotheses. Like the other samples, high CRT is negatively and significantly related to the value-adding device choices ($b = -9.6$%-points at $p = 0.04$). In this sample, we also observe males are less likely to choose the value-based option ($b = -15.8$ at $p = 0.001$). Finally, purchasing experience is negatively associated with value-based device choice meaning that one additional year of purchasing experience is associated with a 1%-point drop ($p = 0.02$) in value-based device choices.

Variables	Model 1c β coefficients (std. errors)	95% CI	Model 2c β coefficients (std. errors)	95% CI
Intercept	0.681*** (0.02)	[0.64, 0.72]	0.77***	[0.32, 1.22]
Incentive	−0.09** (0.04)	[−0.18, 0]	−0.078* (0.04)	[−0.17, 0.01]
Risk sharing	−0.04 (0.04)	[−0.13, 0.05]	−0.055 (0.05)	[−0.15, 0.04]
Evidence	0.074* (0.04)	[−0.01, 0.16]	0.07 (0.05)	[−0.02, 0.16]
Incentive* Risk sharing	0.082 (0.09)	[−0.09, 0.26]	0.104 (0.09)	[−0.07, 0.28]
Incentive* Evidence	−0.169* (0.09)	[−0.34, 0]	−0.189** (0.09)	[−0.36, −0.01]
Risk sharing* Evidence	0.129 (0.09)	[−0.04, 0.3]	0.129 (0.09)	[−0.05, 0.31]
Incentive* Risk sharing* Evidence	0.232 (0.18)	[−0.11, 0.58]	0.344* (0.18)	[−0.02, 0.7]
High CRT			−0.096** (0.05)	[−0.19, 0]
Gender			−0.158*** (0.05)	[−0.25, −0.07]
Procurement experience			−0.011** (0.01)	[−0.02, 0]
Other controls[1]			included not significant (n.s.)	
Sample size	449		449	
R-squared	0.036		0.128	
F-statistic	2.654**		2.485***	

Note(s): *p-value <0.10.; **p-value <0.05.; ***p-value <0.01
[1](organisation size; industry; data collection round; procurement experience; age; procurement position; procurement experience; management experience; LTO; DOSPERT-financial, time)
Source(s): Created by authors

Table 3.
Purchasing managers sample: OLS regressions of choice with heteroskedastic standard errors

7. Discussion and contributions

We have framed our study in agency theory, noting the multilayered agency relationships within the VBP context and the impact of intra- and inter-organisational agency and risk-sharing problems on different decision-makers. We have examined how three factors influence decision-makers to select premium-priced value-adding devices: the impact of cost-saving incentives (in the intra-organisational agency relationships); the impact of risk-sharing contracts (in the inter-organisational agency relationship); and the impact of medical evidence to reduce information asymmetry (in the inter-organisational agency relationship). We have collected data from three different professional samples representing typical decision-makers in hospital purchasing committees: managers, purchasing managers and medical professionals. Table 4 summarises our findings across the three samples.

We find support for the incentive effect in the general manager sample and some support for purchasing manager samples but no support in the medical professional sample (although the effect's direction and magnitude are similar to the other two samples and are as hypothesised). In contrast, our second set of hypotheses regarding the effect of risk-sharing contracts holds for the medical professional sample but not for either managerial sample. In the purchasing manager sample, the direction of the effect is opposite to the hypothesised direction but insignificant. The results for our third set of hypotheses concerning the effect of medical evidence were also mixed. While there was strong support from the medical professional sample, we also observed high failure rates in checks for this manipulation across all three samples. Thus, all the observed effects should be treated with caution.

Hypothesis	Sample	Effect size [95% CI] (p-value)	Interpretation
H1a: Organisational cost-saving incentives decrease managers' likelihood of choosing premium-priced, value-adding products	General managers	−8.3 %-points [−16, −1] ($p = 0.035$)	*Supported*
H1b: Organisational cost-saving incentives decrease medical professionals' likelihood of choosing premium-priced, value-adding products	Medical professionals	−7.8 %-points [−2, 4] ($p = 0.2$)	Not supported
H1c: Organisational cost-saving incentives decrease purchasers' likelihood of choosing premium-priced, value-adding products	Purchasing managers	−7.8 %-points [−17, 1] ($p = 0.08$)	*Weakly supported*
H2a: A risk-sharing contract offered by the supplier increases managers' likelihood of choosing premium-priced, value-adding products	General managers	+5.7 %-points [−2, 13] ($p = 0.14$)	Not supported
H2b: A risk-sharing contract offered by the supplier increases medical professionals' likelihood of choosing premium-priced, value-adding products	Medical professionals	+16 %-points [4, 28] ($p = 0.007$)	*Supported*
H2c: A risk-sharing contract offered by the supplier increases purchasers' likelihood of choosing premium-priced, value-adding products	Purchasing managers	−5 %-points [−15, 4] ($p = 0.23$)	Not supported
H3a: The availability of stronger clinical evidence of the benefits of the value-adding product increases the managers' likelihood of choosing premium-priced, value-adding products	General managers	+3.2 %-points [−4, 11] ($p = 0.4$)	Not supported
H3b: The availability of stronger clinical evidence of the benefits of the value-adding product increases the medical professionals' likelihood of choosing premium-priced, value-adding products	Medical professionals	+15 %-points [3, 27] ($p = 0.01$)	*Supported*
H3c: The availability of stronger clinical evidence of the benefits of the value-adding product increases the purchasers' likelihood of choosing premium-priced, value-adding products	Purchasing managers	+7 %-points [−2, 16] ($p = 0.12$)	Not supported

Source(s): Created by authors

Table 4.
Summary of results

7.1 Theoretical contributions

There have been calls for more multilevel research in supply chain management both theoretically and empirically (Oosterhuis *et al.*, 2005; Carter *et al.*, 2015; Touboulic and Walker, 2015). Management theories provide valuable insights at different levels, so bridging multiple levels of analysis is vital for theoretical development (Harland and Roehrich, 2022). In purchasing contexts in particular, multilevel approaches offer deeper insights into phenomena because purchasing naturally cuts across functional, intra-organisational and inter-organisational domains (Oosterhuis *et al.*, 2005; Carter *et al.*, 2015; Meschnig *et al.*, 2018). Whilst agency theory is a theory already applied at multiple levels, it is predominantly only *one level at a time* that is considered –in purchasing and supply contexts, that level being the inter-organisational one (Matinheikki *et al.*, 2022; Zsidisin, 2022). In contrast, our conceptualisation of VBP as a multilayered agency and risk-sharing problem indicates

that we should not consider agency relationships "one at a time" to align goals between the principal and the agent. Instead, the agency relationships at both intra- and inter-organisational levels must be simultaneously aligned in terms of the incentives and risk sharing for successful employment of more advanced outcome-based contracts, as is the case with value-based procurement of medical devices. As such, our study fully supports Carter *et al.* (2015) in arguing that single-level theorisations can restrict insights into complex phenomena in supply chains.

In prior purchasing studies, there has been a preference for using theories that focus on the organisational viewpoint rather than the behaviour of individual decision-makers (Harland and Roehrich, 2022). While the predominant agency relation when it comes to value-based procurement takes place between buyer and supplier organisation, the individual-level agency relationship may dominate the organisational-level outcomes. Thus, detailed analysis of individual-level behaviour in buyer-supplier agency relationship studies offers the possibility of a deeper understanding of the actual degree of goal incongruence. It may well be that inter-*organisational* goals are well aligned but agency problems emerged through misaligned goals on the *individual* level. We see such multi-tiered agency analysis as particularly pertinent for contexts with high risk in terms of outcomes due to either environmental uncertainty or product/service complexity (and this product complexity may indeed represent a boundary condition for our theoretical contributions).

In contrast to the (purchasing) managers, the incentive condition did not have a significant effect on medical professionals. Likely, such individuals exhibit a stronger tendency to prioritise patient welfare by reducing risks for the patient rather than cost risks for the hospital. In line with agency theory, care professionals play a double agency role of serving patients alongside more cost-conscious hospital administration (Peltokorpi *et al.*, 2020). As such, they exhibit behaviour more in line with assumptions of stewardship rather than agency theory, as the cost incentive did not induce as strong a "self-serving effect" in the decision-making. Administrative pressures on medical professionals to act as economic agents (Vogus *et al.*, 2020) are not always fully congruent with the goal of providing quality care. Therefore, we argue that hospital purchasing organisations should re-evaluate their purchasing incentive schemes more in line with total value contribution (TVC) (Gray *et al.*, 2020) to include value criteria if broader health outcomes are indeed a priority (Fader *et al.*, 2016).

Our study suggests that *inter-organisational risk-sharing contracts* are more appealing to the principal organisation decision-makers who are closer to the operating environments where the risks emerge – in our case, the patient-physician relationship. Medical professionals clearly responded to the treatment, while the other two samples did not. This further supports medical professionals as stewards, since they are more invested in medical outcomes and have expertise in evaluating medical risks such as infections (Abdulsalam *et al.*, 2018). They are, essentially, the agent to the patient, who ultimately bears the burden of any negative medical outcomes, and their risk preferences are thus closely aligned to those of the patient.

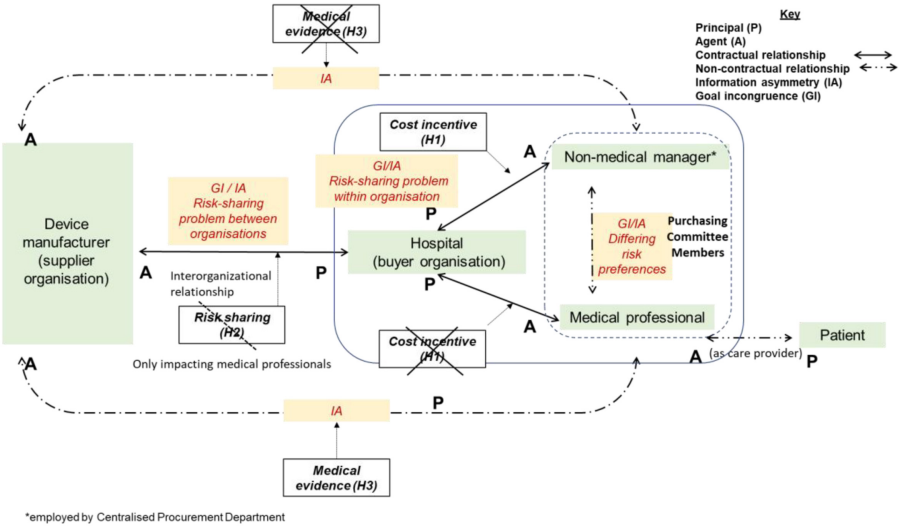
Interestingly, all samples perceived the standard device as riskier than the value-adding device. The differences in the means of device risk perception scores (on a 1–7 Likert scale) for the three samples were (p-value from two-tailed t-test in parentheses): 1.03 (at $p = 0.00$) for general managers, 0.75 (at $p = 0.00$) for medical professors and 0.9 (at $p = 0.00$) for purchasing managers (positive difference indicating that standard device was perceived riskier). Loss aversion is a likely explanation: (purchasing) managers may be willing to take on the risk of increased complications and costs with a sure monetary benefit rather than risk the higher costs with potential gains in *future* costs and medical benefits (that would not benefit their department regardless). Another possible explanation for managers' milder response to the risk-sharing condition is that they were not penalised for risky health outcomes. When not measured and incentivised for total cost avoidance (here in terms of post-operation costs) or

total value contribution, managers as agents of the organisation may simply benignly neglect these aspects (Ellram and Tate, 2021).

Overall, the results regarding the impact of *medical value evidence* were surprising. Not only did we not find statistically significant impacts in the manager and purchasing manager samples, but the manipulation check results were poor compared to other manipulation checks in all samples. The literature indicated that device evidence is used in decision-making and is especially important in the UK for meeting value-based goals (Hurst *et al.*, 2019). Yet, our results suggest a tendency for managers to potentially ignore or misinterpret medical evidence as they may not be trained to evaluate such information. This speaks to Hendry's (2002) alternative explanation to the agency problem, whereby agents are not opportunistic but boundedly rational, and may act against the interest of the principal (here choosing a legitimately proven better value alternative) due to honest incompetence. The medical information was not able to alleviate the (purchasing) managers' information asymmetry as representatives of the buyer organisation, given they were unable to recognise the level of clinical assurance provided. As such, our findings provide support for Antonanzas *et al.* (2019) and Kokshagina and Keränen (2022) who suggest that VBP models necessitate upskilling of procurement personnel. Concerning medical professionals' responses, a possible explanation is that the evidence provided was not differentiated enough compared to their own experiences. Given that evidence in real situations is not standardised, medical professionals may discount evidence in favour of their own best judgement (Kull *et al.*, 2014; Abdulsalam *et al.*, 2018).

An alternative explanation is that managers omit information that does not align with subjective procurement norms. Moses and Åhlström (2008) found that managers in a cross-functional sourcing process often ignore information not in line with their goals. Foerstl *et al.* (2021) also suggest that management decisions are often motivated by subjective social norms. Thus, perhaps rather than not seeking, managers were ignoring the evidence that did not fit with their tasked functional role in purchasing. This supports a need for more training and upskilling of purchasers filling this type of role.

Figure 3 demonstrates the different levels of agency relationships at play in the value-based procurement context in healthcare along with our hypotheses and whether they were



Source(s): Created by authors

Figure 3.
Impact of the hypothesised mechanisms on inter- and intra-organisational agency relationships in medical device purchasing

confirmed with the empirical data (crossed-over hypotheses did not hold). Our results suggest that the individual-level agency relationship goal incongruence, information asymmetry and risk preferences more strongly influence decision-makers choice than the same issues at the inter-organisational level. For (purchasing) managers, cost-incentives had the strongest influence whereas risk-sharing only impacted medical professionals because they are the only ones with something to gain in their intra-organisational/patient agency relationship, and it aligns with their risk preferences (patient risk valued more). Medical evidence only reduced information asymmetry for those decision-makers with a background to be more receptive towards it. Taken together, our results speak to the difficulty of setting the correct mechanisms when intra- and inter-organisational agency relationships are in play in healthcare contexts as well as to the need for more agency research acknowledging the multilevel agency relationships of the studied context.

7.2 Managerial implications

The healthcare sector has been noted as slow to adopt innovations and performance-raising practices, and some of the blame has been attributed to procurement management (Miller *et al.*, 2019; Pedroso *et al.*, 2022). Our study has shed light on three different perspectives in a healthcare purchasing context to understand what impacts the uptake of value-adding medical devices. Understanding the differing preferences of key decision-makers can help solve tensions in multilayered healthcare agency relationships through which medical device procurement often occurs in practice (Atilla *et al.*, 2018).

Our experimental results partially support the claim that cost-based *incentives* in purchasing prevent the selection of value-based options. It also helps explain why VBP appears difficult to implement in practice, especially given the increased outsourcing of healthcare purchasing to professional purchasers. This supports the conclusion that for VBP approaches to be successful, organisations need not just align incentives but redesign criteria where cost is given lower priority relative to value/quality (Gray *et al.*, 2020).

Integrating different preferences in product or service contracting can help change structures and incentives to be more aligned with the total value contribution of the whole organisation. Specifically, we advocate adjusting purchasing incentives (and indeed overall identity, see, e.g. Murfield *et al.*, 2021) away from being heavily cost-focused. This necessitates improved data systems that allow tracking of the total costs and (medical) outcomes of purchasing decisions in healthcare and making these data easily available and interpretable to all stakeholders (Wise *et al.*, 2021). Emphasis on value-added metrics relating to the patient, such as patient-based assessment after a procedure or readmission rates, is critical in enabling the full value potential of purchasing (Gray *et al.*, 2020; Lee *et al.*, 2020).

While conducted from a purchasing perspective, our results also provide insights to suppliers of value-added products or services. Our results demonstrate that (purchasing) managers, who often exert significant control on the tendering criteria, do not appear susceptible to either risk-sharing or enhanced medical evidence of value. This means that suppliers must potentially find more nuanced ways to demonstrate the total cost and value impacts in a language that purchasing understands and is incentivised on. It also suggests purchasing managers may suffer from honest incompetence as agents, which supports the importance for training of purchasing professionals in healthcare.

The managers and procurement managers oversee the supply contract (post-device selection) where the risk-sharing itself would have an impact. Yet, in contrast to the impacts of cost savings, the positive or negative impacts of a risk-sharing contract would only manifest to the hospital as the principal and not be transferred to the employment agency relationship. Hence, the results would suggest that a stronger link should be made between the intra- and inter-organisational agency relationship incentives and risk-sharing

mechanisms to increase the uptake of value-adding devices. Alternatively, outcome-based contracts are being used in complex settings more generally.

7.3 Limitations and future research

Though the ideals that VBP encompasses are not new (e.g. Berwick *et al.*, 2008), there remain barriers to its implementation that require long-term investment. More studies regarding management-stakeholder tensions (e.g. Smith and Lewis, 2011) and the tensions between individual decision-makers and hospital administration are necessary to improve complicated risk, incentive and knowledge structures. We also see benefit in focused research examining effective cross-training for both medical and purchasing staff and the role of teaching hospitals in encouraging the use of value-adding devices to increase VBP practice implementation (e.g. Pedroso *et al.*, 2022).

While our findings provide valuable contributions to both academics and practitioners, the limitations of our study give rise to several potential research avenues. First is the limit on how accurately survey responses will reflect subjects' real-life behaviour (Lonati *et al.*, 2018). Our goal was to make our vignette scenario as realistic as possible for public health procurement, but there are natural limitations to eliciting accurate managerial behaviour. In striving for realism, our vignette also suffers from a lack of wide generalisability since pacemakers speak to issues of life and death when many other procurement decisions are fundamentally less consequential. In addition, managers are not simply motivated by organisational outcomes but also by personal aspirations such as career concerns, which are difficult to mimic in a one-time vignette study. Another limiting factor was being unable to explore whether the same behavioural patterns would be observed if the decision also involved multiple suppliers. Future experiments could be set up more like a traditional public procurement scenario where purchasers are given multiple tenders to evaluate.

An aspect not included in our vignette was the role of product switching, as the choice was between two options from an existing supplier. Many medical devices are considered *physician preference items* where, e.g. medical doctors can have strong brand or type preferences (Robinson, 2008; Montgomery and Schneller, 2007). Hence, we see the risks of product switching (either in terms of medical skill and knowledge levels for the medical professionals) or in terms of supplier switching-related costs (for the purchasing managers) as an interesting avenue to consider regarding VBP uptake.

The low pass rate of the evidence manipulation checks provides a methodological shortcoming preventing us from fully evaluating the evidence effect, but it also opens an intriguing research avenue in the form of more systematic testing of subjects' receptiveness to different forms of (medical) evidence. Research programs examining evidence literacy skills would also be highly relevant in times plagued by increasing levels of disinformation. In addition, only 69% of general managers in the incentive treatment reported after making the decision that they had realised that they could earn real payoffs (see Appendix C, Table C2). For some subjects, the degree of manipulation could have even been stronger if they had realised that they could earn an extra (actual) payment. Though it was clearly stated that these bonus payments were real, participants still interpreted this as hypothetical. Unfortunately, being unable to ascertain specific reasons for this is a limitation of using survey platforms. As such, while it is evident that most subjects experienced some degree of manipulation within the scenario, there is a possibility that our results are biased.

Behaviour may also depend on the funding of health services (private versus public). If more money is spent on pacemakers, does that mean other implantable devices receive the same priority from purchasing managers or are some considered less important? How priority is given based on the perceived item value is an aspect warranting further research.

We did note an interesting result that women were more likely to select the value-adding device across treatment groups. This follows the expected behaviour in economics where women tend to behave more risk aversely, even if the risk is comparatively low (Weber *et al.*, 2002). While the NHS is female-dominated, men are overrepresented in senior management roles at 53% (NHS Employers, 2019), where they are more likely to have an influence on VBP. As such, the role of women in enacting change in healthcare policy opens another interesting avenue for future research.

Our results taken together echo the recent calls for increased data collection and data sharing regarding long-term value outcomes in healthcare (Antonanzas *et al.*, 2019; Kokshagina and Keränen, 2022). Pedroso *et al.* (2022) concluded in their Danish study that value-based procurement initiatives succeeded only when clinical data were available for several years. Tying such improved health outcome data better to the personal incentives of purchasing committee members could lead to increased uptake of value-adding devices. A shift from cost-based to more hybrid incentive structures, where part of the bonus is determined by past medical outcomes of the device categories that the employee is responsible for, could lead to different purchase selections in the long term. It could also offer an interesting future experimental setup to evaluate the most effective combinations of incentives.

Notes

1. We thank an anonymous reviewer for their advice and insights considering the risk-sharing problem and its impact on our study context.
2. For product details of the envelope see, e.g. https://www.medtronic.com/us-en/healthcare-professionals/products/cardiac-rhythm/absorbable-antibacterial-envelopes/tyrx-envelope.html (Accessed 03/2022). The research was neither financially supported nor influenced by Medtronic but the device acted simply as an inspiration for the study.
3. See, e.g. Carlson *et al.* (2017) Medtronic's TYRX envelopes aimed at reducing heart device infections. *Star Tribune* June 10. Available at (accessed 03/2022): https://www.startribune.com/medtronic-s-tyrx-envelopes-aimed-at-reducing-heart-device-infections/427535213/
4. See, e.g. Oxford Centre for Evidence-based Medicine. https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence. It should be noted that sometimes a meta-analysis of RCTs is considered the strongest form of evidence.
5. An aggregate category, which represents miscellaneous healthcare professions providing a wide range of diagnostic, technical, therapeutic and support services in connection with health care (see, e.g. Wikipedia: "Allied Health Professions").

References

Abdulsalam, Y. and Schneller, E. (2019), "Hospital supply expenses: an important ingredient in health services research", *Medical Care Research and Review*, Vol. 76 No. 2, pp. 240-252, doi: 10.1177/1077558717719928.

Abdulsalam, Y., Gopalakrishnan, M., Maltz, A. and Schneller, E. (2018), "The impact of physician-hospital integration on hospital supply management", *Journal of Operations Management*, Vol. 57 No. 1, pp. 11-22, doi: 10.1016/j.jom.2018.01.001.

Ahmed, F.Z., Fullwood, C., Zaman, M., Qamruddin, A., Cunnington, C., Mamas, M.A., Sandoe, J., Motwani, M. and Zaidi, A. (2019), "Cardiac implantable electronic device (CIED) infections are expensive and associated with prolonged hospitalisation: UK Retrospective Observational Study", *PLoS One*, Vol. 14 No. 1, e0206611, doi: 10.1371/journal.pone.0206611.

Albarqouni, L., Hoffmann, T., Straus, S., Olsen, N.R., Young, T., Ilic, D., Shaneyfelt, T., Haynes, R.B., Guyatt, G. and Glasziou, P. (2018), "Core competencies in evidence-based practice for health professionals: consensus statement based on a systematic review and Delphi survey", *JAMA Network Open*, Vol. 1 No. 2, e180281, doi: 10.1001/jamanetworkopen.2018.0281.

Anderson, J.C. and Wynstra, F. (2010), "Purchasing higher-value, higher-price offerings in business markets", *Journal of Business-to-Business Marketing*, Vol. 17 No. 1, pp. 29-61, doi: 10.1080/10517120903000363.

Anderson, J.C., Thomson, J.B. and Wynstra, F. (2000), "Combining value and price to make purchase decisions in business markets", *International Journal of Research in Marketing*, Vol. 17 No. 4, pp. 307-329, doi: 10.1016/s0167-8116(00)00029-x.

Antonanzas, F., Juárez-Castelló, C., Lorente, R. and Rodríguez-Ibeas, R. (2019), "The use of risk-sharing contracts in healthcare: theoretical and empirical assessments", *Pharmacoeconomics*, Vol. 37 No. 12, pp. 1469-1483, doi: 10.1007/s40273-019-00838-w.

Aronow, P.M., Baron, J. and Pinson, L. (2019), "A note on dropping experimental subjects who fail a manipulation check", *Political Analysis*, Vol. 27 No. 4, pp. 572-589, doi: 10.1017/pan.2019.5.

Atilla, E.A., Steward, M., Wu, Z. and Hartley, J.L. (2018), "Triadic relationships in healthcare", *Business Horizons*, Vol. 61 No. 2, pp. 221-228, doi: 10.1016/j.bushor.2017.11.004.

Bae, J.M. (2015), "Value-based medicine: concepts and application", *Epidemiology and Health*, Vol. 37, e2015014, doi: 10.4178/epih/e2015014.

Ball, G.P., Shah, R. and Donohue, K. (2018), "The decision to recall: a behavioral investigation in the medical device industry", *Journal of Operations Management*, Vol. 62, pp. 1-15, doi: 10.1016/j.jom.2018.07.003.

Bearden, W.O., Money, R.B. and Nevins, J.L. (2006), "A measure of long-term orientation: development and validation", *Journal of the Academy of Marketing Science*, Vol. 34 No. 3, pp. 456-467, doi: 10.1177/0092070306286706.

Bech, M. and Gyrd-Hansen, D. (2005), "Effects coding in discrete choice experiments", *Health Economics*, Vol. 14 No. 10, pp. 1079-1083, doi: 10.1002/hec.984.

Berwick, D.M., Nolan, T.W. and Whittington, J. (2008), "The triple aim: care, health, and cost", *Health Affairs*, Vol. 27 No. 3, pp. 759-769, doi: 10.1377/hlthaff.27.3.759.

Blais, A.R. and Weber, E.U. (2006), "A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations", *Judgment and Decision Making*, Vol. 1 No. 1, pp. 33-47, doi: 10.1017/s1930297500000334.

Carlson, J.J., Chen, S. and Garrison, L.P. (2017), "Performance-based risk-sharing arrangements: an updated international review", *Pharmacoeconomics*, Vol. 35 No. 10, pp. 1063-1072, doi: 10.1007/s40273-017-0535-z.

Carter, C.R., Meschnig, G. and Kaufmann, L. (2015), "Moving to the next level: why our discipline needs more multilevel theorization", *Journal of Supply Chain Management*, Vol. 51 No. 4, pp. 94-102, doi: 10.1111/jscm.12083.

Chae, S., Lawson, B., Kull, T. and Choi, T. (2019), "To insource or outsource the sourcing? A behavioral investigation of the multi-tier sourcing decision", *International Journal of Operations and Production Management*, Vol. 39 No. 3, pp. 385-405, doi: 10.1108/ijopm-04-2018-0231.

Conrad, D.A. (2015), "The theory of value-based payment incentives and their application to health care", *Health Services Research*, Vol. 50 No. S2, pp. 2057-2089, doi: 10.1111/1475-6773.12408.

Cook, J.P., Vernon, J.A. and Manning, R. (2008), "Pharmaceutical risk-sharing agreements", *Pharmacoeconomics*, Vol. 26 No. 7, pp. 551-556, doi: 10.2165/00019053-200826070-00002.

Davis, J.H., Schoorman, F.D. and Donaldson, L. (1997), "Toward a stewardship theory of management", *Academy of Management Review*, Vol. 22 No. 1, pp. 20-47, doi: 10.5465/amr.1997.9707180258.

Dumond, E.J. (1996), "Applying value-based management to procurement", *International Journal of Physical Distribution and Logistics Management*, Vol. 26 No. 1, pp. 5-24, doi: 10.1108/09600039610108566.

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C.R. and Kaufmann, L. (2021), "On making experimental design choices: discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes", *Journal of Operations Management*, Vol. 67 No. 2, pp. 261-275, doi: 10.1002/joom.1128.

Ellram, L.M. and Tate, W.L. (2021), "Cost avoidance: not everything that counts is counted", *Journal of Business Logistics*, Vol. 42 No. 4, pp. 406-427, doi: 10.1111/jbl.12262.

European Commission (2021), "Expert panel on effective ways of investing in health (EXPH), opinion on public procurement in healthcare systems", Publications Office of the European Union, Luxembourg, available at: https://ec.europa.eu/health/sites/default/files/expert_panel/docs/027_public_proc_healthcare_sys_en.pdf

Fader, A.N., Xu, T., Dunkin, B.J. and Makary, M.A. (2016), "A perverse quality incentive in surgery: implications of reimbursing surgeons less for doing laparoscopic surgery", *Surgical Endoscopy*, Vol. 30 No. 11, pp. 4665-4667, doi: 10.1007/s00464-016-5170-9.

Foerstl, K., Franke, H. and Cataldo, Z. (2021), "What drives managers to insource production? Evidence from a behavioural experiment", *Journal of Purchasing and Supply Management*, Vol. 27 No. 4, 100715, doi: 10.1016/j.pursup.2021.100715.

Formentini, M. and Romano, P. (2016), "Towards supply chain collaboration in B2B pricing: a critical literature review and research agenda", *International Journal of Operations and Production Management*, Vol. 36 No. 7, pp. 734-756, doi: 10.1108/ijopm-03-2015-0124.

Frederick, S. (2005), "Cognitive reflection and decision making", *Journal of Economic Perspectives*, Vol. 19 No. 4, pp. 25-42, doi: 10.1257/089533005775196732.

Fuloria, P.C. and Zenios, S.A. (2001), "Outcomes-adjusted reimbursement in a health-care delivery system", *Management Science*, Vol. 47 No. 6, pp. 735-751, doi: 10.1287/mnsc.47.6.735.9816.

Ganzach, Y., Ellis, S., Pazy, A. and Ricci-Siag, T. (2008), "On the perception and operationalization of risk perception", *Judgment and Decision Making Journal*, Vol. 3 No. 4, pp. 317-324, doi: 10.1017/s1930297500000887.

Gardner, J.W., Boyer, K.K. and Gray, J.V. (2015), "Operational and strategic information processing: complementing healthcare IT infrastructure", *Journal of Operations Management*, Vol. 33 No. 1, pp. 123-139, doi: 10.1016/j.jom.2014.11.003.

Gomila, R. (2021), "Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis", *Journal of Experimental Psychology: General*, Vol. 150 No. 4, pp. 700-709, doi: 10.1037/xge0000920.

Gottschalg, O. and Zollo, M. (2007), "Interest alignment and competitive advantage", *Academy of Management Review*, Vol. 32 No. 2, pp. 418-437.

Gray, J.V., Helper, S. and Osborn, B. (2020), "Value first, cost later: total value contribution as a new approach to sourcing decisions", *Journal of Operations Management*, Vol. 66 No. 6, pp. 735-750, doi: 10.1002/joom.1113.

Grundy, Q. (2016), "'Whether something cool is good enough': the role of evidence, sales representatives and nurses' expertise in hospital purchasing decisions", *Social Science and Medicine*, Vol. 165, pp. 82-91, doi: 10.1016/j.socscimed.2016.07.042.

Harland, C.M. and Roehrich, J.K. (2022), "Systems levels in purchasing and supply chain management (PSCM) research: exploring established and novel theories to address PSCM problems and challenges", in *Handbook of Theories for Purchasing, Supply Chain and Management Research*, Edward Elgar Publishing, pp. 63-79.

Helper, S., Gray, J.V., Hughes, M.M. and Roman, A.V. (2021), "Public policy and operations management", *Journal of Operations Management*, Vol. 67 No. 7, pp. 780-802, doi: 10.1002/joom.1160.

Hinrichs-Krapels, S., Ditewig, B., Boulding, H., Chalkidou, A., Erskine, J. and Shokraneh, F. (2022), "Purchasing high-cost medical devices and equipment in hospitals: a systematic review", *BMJ Open*, Vol. 12 No. 9, e057516, doi: 10.1136/bmjopen-2021-057516.

Hofmann, E., Maucher, D., Kotula, M. and Kreienbrink, O. (2014), *Performance Measurement and Incentive Systems in Purchasing: More than Just Savings*, Springer Science and Business Media, Berlin Heidelberg, Vol. 3, p. 182.

Hurst, L., Mahtani, K., Pluddemann, A., Lewis, S., Harvey, K., Briggs, A., ... Heneghan, C. (2019), "Defining value-based healthcare in the NHS", Centre for Evidence-Based Medicine, Report 2019/04.

Jensen, M.C. and Heckling, W.H. (1995), "Specific and general knowledge, and organizational structure", *Journal of Applied Corporate Finance*, Vol. 8 No. 2, pp. 4-18.

Kaczmarek, S. (2017), "Rethinking board diversity with the behavioural theory of corporate governance: opportunities and challenges for advances in theorising", *Journal of Management and Governance*, Vol. 21 No. 4, pp. 879-906, doi: 10.1007/s10997-016-9362-4.

Kähkönen, A.K. and Lintukangas, K. (2018), "Key dimensions of value creation ability of supply management", *International Journal of Operations and Production Management*, Vol. 38 No. 4, pp. 979-996, doi: 10.1108/ijopm-04-2016-0189.

Kauppi, K. and Van Raaij, E.M. (2015), "Opportunism and honest incompetence—seeking explanations for noncompliance in public procurement", *Journal of Public Administration Research and Theory*, Vol. 25 No. 3, pp. 953-979, doi: 10.1093/jopart/mut081.

Kienzler, M. (2018), "Value-based pricing and cognitive biases: an overview for business markets", *Industrial Marketing Management*, Vol. 68, pp. 86-94, doi: 10.1016/j.indmarman.2017.09.028.

Kirchoff, J.F., Omar, A. and Fugate, B.S. (2016), "A behavioral theory of sustainable supply chain management decision making in non-exemplar firms", *Journal of Supply Chain Management*, Vol. 52 No. 1, pp. 41-65, doi: 10.1111/jscm.12098.

Kokshagina, O. and Keränen, J. (2022), "Institutionalizing value-based healthcare in a service system: a policy and document analysis over three decades", *Journal of Business and Industrial Marketing*, Vol. 37 No. 8, pp. 1607-1622, doi: 10.1108/jbim-08-2020-0380.

Kull, T.J., Oke, A. and Dooley, K.J. (2014), "Supplier selection behavior under uncertainty: contextual and cognitive effects on risk perception and choice", *Decision Sciences*, Vol. 45 No. 3, pp. 467-505, doi: 10.1111/deci.12078.

Lee, S.J., Venkataraman, S., Heim, G.R., Roth, A.V. and Chilingerian, J. (2020), "Impact of the value-based purchasing program on hospital operations outcomes: an econometric analysis", *Journal of Operations Management*, Vol. 66 Nos 1-2, pp. 151-175, doi: 10.1002/joom.1057.

Lonati, S., Quiroga, B.F., Zehnder, C. and Antonakis, J. (2018), "On doing relevant and rigorous experiments: review and recommendations", *Journal of Operations Management*, Vol. 64 No. 1, pp. 19-40, doi: 10.1016/j.jom.2018.10.003.

Lonsdale, C. and Watson, G. (2005), "The internal client relationship, demand management and value for money: a conceptual model", *Journal of Purchasing and Supply Management*, Vol. 11 No. 4, pp. 159-171, doi: 10.1016/j.pursup.2005.11.001.

Matinheikki, J., Kauppi, K., Brandon–Jones, A. and van Raaij, E.M. (2022), "Making agency theory work for supply chain relationships: a systematic review across four disciplines", *International Journal of Operations and Production Management*, Vol. 42 No. 13, pp. 299-334, doi: 10.1108/ijopm-12-2021-0757.

McKone-Sweet, K.E., Hamilton, P. and Willis, S.B. (2005), "The ailing healthcare supply chain: a prescription for change", *Journal of Supply Chain Management*, Vol. 41 No. 1, pp. 4-17, doi: 10.1111/j.1745-493x.2005.tb00180.x.

Meehan, J., Menzies, L. and Michaelides, R. (2017), "The long shadow of public policy; Barriers to a value-based approach in healthcare procurement", *Journal of Purchasing and Supply Management*, Vol. 23 No. 4, pp. 229-241, doi: 10.1016/j.pursup.2017.05.003.

Meschnig, G., Carter, C. and Kaufmann, L. (2018), "Conducting multilevel studies in purchasing and supply management research", *Journal of Purchasing and Supply Management*, Vol. 24 No. 4, pp. 338-342, doi: 10.1016/j.pursup.2017.11.001.

Miller, F.A., Lehoux, P., Peacock, S., Rac, V.E., Neukomm, J., Barg, C., Bytautas, J.P. and Krahn, M. (2019), "How procurement judges the value of medical technologies: a review of healthcare tenders", *International Journal of Technology Assessment in Health Care*, Vol. 35 No. 1, pp. 50-55, doi: 10.1017/s0266462318003756.

Montgomery, K. and Schneller, E.S. (2007), "Hospitals' strategies for orchestrating selection of physician preference items", *The Milbank Quarterly*, Vol. 85 No. 2, pp. 307-335, doi: 10.1111/j.1468-0009.2007.00489.x.

Moses, A. and Åhlström, P. (2008), "Problems in cross-functional sourcing decision processes", *Journal of Purchasing and Supply Management*, Vol. 14 No. 2, pp. 87-99, doi: 10.1016/j.pursup.2007.11.003.

Murfield, M.L.U., Ellram, L.M. and Giunipero, L.C. (2021), "Moving purchasing and supply management beyond a cost-focused identity", *Journal of Purchasing and Supply Management*, Vol. 27 No. 3, 100687, doi: 10.1016/j.pursup.2021.100687.

NHS Employers (2019), "Gender in the NHS Infographic", Gender in the NHS infographic | NHS Employers, Retrieved 21 March 2023.

Obremskey, W.T., Dail, T. and Jahangir, A.A. (2012), "Value-based purchasing of medical devices", *Clinical Orthopaedics and Related Research*, Vol. 470 No. 4, pp. 1054-1064, doi: 10.1007/s11999-011-2147-9.

Oosterhuis, M., Molleman, E. and van der Vaart, T. (2005), "Multilevel issues in supply chain management", in *Research Methodologies in Supply Chain Management: In Collaboration with Magnus Westhaus*, pp. 283-297.

Pedroso, C.B., Beaulieu, M., Allerup, L.D. and Rebolledo, C. (2022), "Fostering innovation through procurement in the healthcare sector: the Danish experience", *Healthcare Quarterly*, Vol. 24 No. 4, pp. 22-26, doi: 10.12927/hcq.2022.26715.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z. and Damer, E. (2022), "Data quality of platforms and panels for online behavioral research", *Behavior Research Methods*, Vol. 54, pp. 1643-1662.

Peltokorpi, A., Matinheikki, J., Lehtinen, J. and Rajala, R. (2020), "Revisiting the unholy alliance of health-care operations: payor–provider integration of occupational health services", *International Journal of Operations and Production Management*, Vol. 40 No. 4, pp. 357-387, doi: 10.1108/ijopm-04-2019-0326.

Prada, G. (2016), "Value-based procurement: Canada's healthcare imperative", *Healthcare Management Forum*, Vol. 29 No. 4, pp. 162-164, doi: 10.1177/0840470416646119.

Riedl, D.F., Kaufmann, L., Zimmermann, C. and Perols, J.L. (2013), "Reducing uncertainty in supplier selection decisions: antecedents and outcomes of procedural rationality", *Journal of Operations Management*, Vol. 31 Nos 1-2, pp. 24-36, doi: 10.1016/j.jom.2012.10.003.

Robinson, J.C. (2008), "Value-based purchasing for medical devices", *Health Affairs*, Vol. 27 No. 6, pp. 1523-1531, doi: 10.1377/hlthaff.27.6.1523.

Rungtusanatham, M., Wallin, C. and Eckerd, S. (2011), "The vignette in a scenario-based role-playing experiment", *Journal of Supply Chain Management*, Vol. 47 No. 3, pp. 9-16, doi: 10.1111/j.1745-493x.2011.03232.x.

Sanderson, J., Lonsdale, C. and Mannion, R. (2019), "What's needed to develop strategic purchasing in healthcare? Policy lessons from a realist review", *International Journal of Health Policy and Management*, Vol. 8 No. 1, pp. 4-17, doi: 10.15171/ijhpm.2018.93.

Smith, W.K. and Lewis, M.W. (2011), "Toward a theory of paradox: a dynamic equilibrium model of organizing", *Academy of Management Review*, Vol. 36 No. 2, pp. 381-403, doi: 10.5465/amr.2011.59330958.

Sorenson, C. and Kanavos, P. (2011), "Medical technology procurement in Europe: a cross-country comparison of current practice and policy", *Health Policy*, Vol. 100 No. 1, pp. 43-50, doi: 10.1016/j.healthpol.2010.08.001.

Tarakji, K.G., Mittal, S., Kennergren, C., Corey, R., Poole, J.E., Schloss, E., Gallastegui, J., Pickett, R.A., Evonich, R., Philippon, F., Roark, S.F., Sorrentino, D., Sholevar, D., Cronin, E., Berman, B., Riggio, D., Biffi, M., Khan, H., Silver, M.T., Collier, J., Eldadah, Z., Wright, D.J., Lande, J.D., Lexcen, D.R., Cheng, A., Wilkoff, B.L. and McComb, J.M. (2019), "Antibacterial envelope to prevent cardiac implantable device infection", *New England Journal of Medicine*, Vol. 380 No. 20, pp. 1895-1905, doi: 10.1056/nejmoa1901111.

Toplak, M.E., West, R.F. and Stanovich, K.E. (2011), "The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks", *Memory and Cognition*, Vol. 39 No. 7, pp. 1275-1289, doi: 10.3758/s13421-011-0104-1.

Tosi, H.L., Katz, J.P. and Gomez-Mejia, L.R. (1997), "Disaggregating the agency contract: the effects of monitoring, incentive alignment, and term in office on agent decision making", *Academy of Management Journal*, Vol. 40 No. 3, pp. 584-602.

Touboulic, A. and Walker, H. (2015), "Theories in sustainable supply chain management: a structured literature review", *International Journal of Physical Distribution and Logistics Management*, Vol. 45 Nos 1/2, pp. 16-42, doi: 10.1108/ijpdlm-05-2013-0106.

Töytäri, P., Keränen, J. and Rajala, R. (2017), "Barriers to implementing value-based pricing in industrial markets: a micro-foundations perspective", *Journal of Business Research*, Vol. 76, pp. 237-246, doi: 10.1016/j.jbusres.2016.04.183.

Tse, Y.K., Zhang, M. and Jia, F. (2018), "The effects of risk and reward sharing on quality performance", *International Journal of Operations and Production Management*, Vol. 38 No. 12, pp. 2367-2388, doi: 10.1108/ijopm-12-2016-0759.

Villena, V.H., Lu, G., Gomez-Mejia, L.R. and Revilla, E. (2018), "Is top management team-supply chain manager interaction the missing link? An analysis of risk-bearing antecedents for supply chain managers", *International Journal of Operations and Production Management*, Vol. 38 No. 8, pp. 1640-1663, doi: 10.1108/ijopm-05-2017-0258.

Vogus, T.J., Gallan, A., Rathert, C., El-Manstrly, D. and Strong, A. (2020), "Whose experience is it anyway? Toward a constructive engagement of tensions in patient-centered health care", *Journal of Service Management*, Vol. 31 No. 5, pp. 979-1013, doi: 10.1108/josm-04-2020-0095.

Wang, Y.A., Sparks, J., Gonzales, J.E., Hess, Y.D. and Ledgerwood, A. (2017), "Using independent covariates in experimental designs: quantifying the trade-off between power boost and Type I error inflation", *Journal of Experimental Social Psychology*, Vol. 72, pp. 118-124, doi: 10.1016/j.jesp.2017.04.011.

Weber, E.U., Blais, A.R. and Betz, N.E. (2002), "A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors", *Journal of Behavioral Decision Making*, Vol. 15 No. 4, pp. 263-290, doi: 10.1002/bdm.414.

Wenzl, M. and Mossialos, E. (2018), "Prices for cardiac implant devices may be up to six times higher in the US than in some European countries", *Health Affairs*, Vol. 37 No. 10, pp. 1570-1577, doi: 10.1377/hlthaff.2017.1367.

Wiens, S. and Nilsson, M.E. (2017), "Performing contrast analysis in factorial designs: from NHST to confidence intervals and beyond", *Educational and Psychological Measurement*, Vol. 77 No. 4, pp. 690-715, doi: 10.1177/0013164416668950.

Wise, S., Hall, J., Haywood, P., Khana, N., Hossain, L. and van Gool, K. (2021), "Paying for value: options for value-based payment reform in Australia", *Australian Health Review*, Vol. 46 No. 2, pp. 129-133, doi: 10.1071/ah21115.

Wong, R.S. (2023), "An experimental investigation of attribute framing effects on risky sourcing behaviour: the mediating role of attention allocated to suppliers' quality information", *International Journal of Operations and Production Management*, Vol. 43 No. 13, pp. 205-225, doi: 10.1108/ijopm-09-2022-0604.

Yan, T., Ribbink, D. and Pun, H. (2018), "Incentivizing supplier participation in buyer innovation: experimental evidence of non-optimal contractual behaviors", *Journal of Operations Management*, Vol. 57 No. 1, pp. 36-53, doi: 10.1016/j.jom.2017.12.001.

Zsidisin, G.A. (2022), "13. Agency theory in purchasing and supply management", in *Handbook of Theories for Purchasing, Supply Chain and Management Research*, p. 186.

Zu, X. and Kaynak, H. (2012), "An agency theory perspective on supply chain quality management", *International Journal of Operations and Production Management*, Vol. 32 No. 4, pp. 423-446, doi: 10.1108/01443571211223086.

Appendix

The supplementary material for this article can be found online.

Corresponding author
Katie Kenny can be contacted at: katie.kenny@aalto.fi