

Pricing models for diagnostic AI based on qualitative insights from healthcare decision makers

Received: 3 October 2025

Accepted: 19 February 2026

Cite this article as: Kirchhoff, J., Berns, F., Schieder, C. *et al.* Pricing models for diagnostic AI based on qualitative insights from healthcare decision makers. *npj Digit. Med.* (2026). <https://doi.org/10.1038/s41746-026-02501-z>

Jan Kirchhoff, Fabian Berns, Christian Schieder & Johannes Schobel

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Pricing Models for Diagnostic AI Based on Qualitative Insights from Healthcare Decision Makers

Jan Kirchhoff^{1,2*}, Fabian Berns², Christian Schieder³,
Johannes Schobel¹

¹DigiHealth Institute, Neu-Ulm University of Applied Sciences,
Wileystraße 1, Neu-Ulm, 89231, Germany.

²medicalvalues GmbH, Haid-und-Neu-Straße 18, Karlsruhe, 76131,
Germany.

³Weiden Business School, OTH Amberg-Weiden, Hetzenrichter Weg 15,
Weiden, 92637, Germany.

*Corresponding author(s). E-mail(s): jan.kirchhoff@medicalvalues.de;
Contributing authors: fabian.berns@medicalvalues.de;
c.schieder@oth-aw.de; johannes.schobel@hnu.de;

Abstract

AI-enabled diagnostic decision support systems (DDSS) could improve diagnostic accuracy and efficiency, yet adoption is often impeded by pricing approaches that rely on opaque technical usage metrics. We examined how pricing can remain clinically legible and budgetable while accounting for AI-specific technical and organizational cost drivers. We conducted semi-structured interviews with healthcare decision makers (n=17) across hospital, outpatient, laboratory, and industry settings and conducted a deductive-inductive thematic analysis. Ten themes emerged, including widespread resistance to purely usage-based pricing and strong preferences for transparency and predictability. Participants supported hybrid models combining a base fee with variable components defined in clinically meaningful units (per patient, per test, or per episode) and emphasized reimbursement alignment alongside integration, training, and support as integral value elements. Outcome-linked payment was viewed as ethically compelling but operationally difficult. We synthesize these findings into stakeholder-informed design principles and actionable recommendations for pricing models that facilitate procurement, reimbursement fit, and sustainable scaling of diagnostic AI.

Introduction

Artificial intelligence (AI) is increasingly shaping many aspects of daily life, from communication and education to industrial production and scientific discovery. Healthcare is increasingly adopting AI applications across the continuum of care, promising improvements in efficiency, accuracy, and patient outcomes [1]. Within medicine, clinical diagnostics is a particularly promising field. Diagnostic Decision Support Systems (DDSS) process heterogeneous clinical data, laboratory results, patient history, images, and unstructured notes to generate differential diagnoses, risk assessments, and guidelines-based suggestions [2].

Although regulatory and reimbursement examples here are drawn primarily from Germany, Austria, and Switzerland, the underlying challenges mirror those of other high-income systems (United States, United Kingdom, Japan). As workloads grow, populations age, and specialization increases, the demand for AI-enabled support tools increases [3]. DDSS can improve diagnostic accuracy, reduce treatment time, and support adherence to complex guidelines [4]. However, pricing and reimbursement mechanisms have not evolved at the same rate. Many systems rely on flat licenses or rudimentary usage metrics that do not reflect ongoing AI costs [5], making financing a primary barrier to scaling when suitable payment routes are absent [6]. Key cost drivers, including runtime inference, continuous model updates, GPU-based infrastructure, and emerging compliance obligations such as the EU AI Act, introduce recurring usage-dependent expenses that are not captured by current models [7]. A related limitation is the frequent lack of differentiation between conventional software functionality and AI-specific capabilities. AI systems require sustained training, real-time inference, and performance monitoring; architecture choices result in different infrastructure demands [8]. Embedded or indirect use (for example, background processes) is poorly captured by per-user licensing and can deter investment in scalable or domain-specific solutions, especially for rare diseases with limited volume [9]. Inference itself is a major cost driver, while cloud deployment adds variability and shifts spending toward operational expenditure [10, 11].

At the vendor level, AI shifts software cost structures toward recurring operational expenditure, faster innovation cycles, and rapid feature replication (often aided by generative AI), which can erode the fit of traditional seat-based or volume-licensing models [12]. Therefore, competitive advantage shifts to integration quality, regulatory compliance, and explainability, which are prerequisites for trust and market access in clinical care [13]. The growing use of open-source AI can reduce license fees, but transfers validation, deployment, and maintenance burdens to providers, motivating hybrid strategies that pair free model access with paid integration and support [12]. Indirect or embedded use also complicates cost attribution and suggests alternatives to user licenses, such as pricing by transactions, processed data points, or clinical episodes [14]. As systems move toward value-based care, payment should increasingly reflect outcomes, which requires robust impact measurement and reimbursement alignment [15]. Related debates on value-based pricing and health technology assessment (HTA) for digital and AI-enabled tools focus on linking reimbursement to demonstrated clinical benefit and cost-effectiveness [14, 17, 19]. Simultaneously, artificial intelligence

systems entail lifecycle costs encompassing phases of development, validation, deployment, monitoring, and updating. These dynamics pose challenges to static licensing frameworks and inspire the formulation of novel methodologies for the allocation of costs and value [7, 20].

Recent work begins to translate these considerations into explicit reimbursement logic for diagnostic AI. Conceptual and modeling studies have outlined how AI services might be integrated into existing fee schedules, for example, by defining AI-specific contribution levels, mapping them to established point systems such as GOÄ, EBM, or CPT, and evaluating their impact in clinical scenarios [5, 6, 18, 28]. Such frameworks demonstrate that pricing models can be designed to be auditable and compatible with learning healthcare systems, and highlight typical trade-offs between technical usage metrics, clinical value, and budget constraints.

Despite this progress, the existing literature offers limited empirical insight into how pricing models are perceived by healthcare organizations that must procure, budget, and operationalize diagnostic AI. Most contributions focus on conceptual pricing schemes, reimbursement policy options, or health-economic modeling, while the perspectives of clinicians, laboratory leaders, and administrative decision-makers on acceptable and sustainable AI pricing remain underexplored [16, 22, 28]. In particular, there is little evidence on which design features (e.g., billable units, hybrid structures, integration and service components, or outcome-linked elements) are viewed as compatible with clinical workflows and institutional budget logics.

As a research gap, there is insufficient empirical evidence on how healthcare decision makers evaluate, compare, and ultimately accept AI pricing architectures that must simultaneously (i) fit procurement and reimbursement logic and (ii) reflect AI-specific technical and organizational cost drivers in a way that remains clinically legible and budgetable.

Accordingly, this article addresses the following research question: *How should pricing models for AI-enabled diagnostic decision support be structured so that they remain understandable and budgetable for procurement and reimbursement, yet still capture AI-specific technical and organizational cost drivers and usage patterns?* To answer this question, we conducted semi-structured interviews with clinical decision makers in inpatient and outpatient care and laboratory settings and performed a thematic analysis to derive a stakeholder-informed framework for future pricing strategies that reflects technical, clinical and economic dimensions. This work translates these insights into a set of stakeholder-informed design principles for AI pricing, designed to support adoption while maintaining fairness and scalability. The Results section synthesizes the interview findings into ten themes; the Discussion section derives implications for pricing models, reimbursement frameworks, and future research; and the Methods section describes the qualitative design and analysis. By translating technical cost drivers into clinically meaningful units and defining policy-compatible models, the findings offer actionable input for reimbursement design, procurement guidance, and regulation of AI-enabled diagnostic support.

Results

The interviews yielded a broad spectrum of perspectives on economic models for AI-enabled diagnostic decision support systems (DDSS). Rather than focusing on isolated accounts, the analysis examines patterns between participants, highlighting shared priorities as well as points of divergence. Ten themes were identified and subsequently grouped into four overarching domains to improve conceptual coherence and to link themes to pricing design principles: (1) *Pricing Architecture and Budget Predictability*, capturing preferences for clinically legible, transparent, and forecastable pricing structures; (2) *Reimbursement and Value Evidence*, reflecting expectations regarding reimbursement alignment and the feasibility of outcome-linked elements; (3) *Implementation Fit and Market Segmentation*, emphasizing integration quality, procurement simplicity, and heterogeneity across provider types; and (4) *Lifecycle Services and Equity*, addressing service expectations, long-term support, and access-related concerns. Each theme is reported individually, but the domain structure clarifies how these findings interrelate and how they collectively motivate stakeholder-informed design principles. For transparency, we provide simple prevalence counts to show how many participants explicitly mentioned each theme, without interpreting these numbers as comparative indicators of importance.

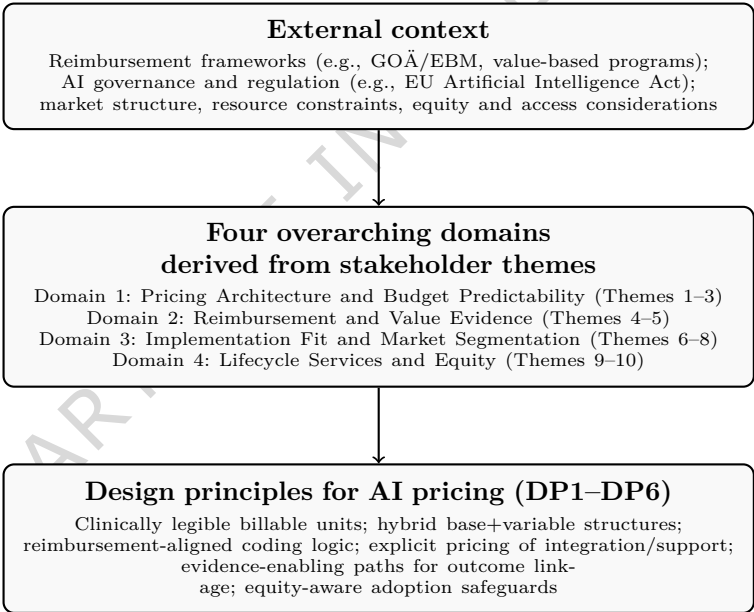


Fig. 1 Conceptual summary of the stakeholder-informed framework for pricing AI-enabled diagnostic decision support systems. External contextual factors (top) shape four overarching domains derived from the interview themes (middle), which in turn converge into a set of design principles for clinically meaningful, predictable, and technically faithful pricing models (bottom).

Emerging themes from the analysis

Figure 1 summarizes how the ten themes and four overarching domains relate to the wider contextual factors and the resulting design principles for AI pricing. It visualizes the hierarchical structure of influences, beginning with regulatory and reimbursement conditions and cascading down toward stakeholder-derived priorities. This representation also clarifies how individual themes collectively shape coherent pricing principles, rather than operating as isolated considerations.

Domain 1: Pricing Architecture and Budget Predictability

Themes 1 to 3 fall into this domain, capturing the emphasis of participants on transparent, fair, and predictable pricing structures that align with clinical budgeting.

Theme 1: Resistance to pure token- or pay-per-use models. Most participants ($n = 13/17$, 76%) considered pricing schemes based solely on technical usage metrics, such as token counts, API calls, or computation time, to be economically and operationally misaligned with healthcare practice. Although such measures may be meaningful for technology vendors or cloud providers, they were described by interviewees as too abstract and opaque for hospitals and laboratories, where budgets are typically planned in clinically tangible units such as per patient, per test, or per case. Without a clear mapping to these established categories, participants emphasized, usage-based models are unlikely to gain acceptance. This concern also contributed to a wider consensus: a strong demand for pricing structures that are transparent, predictable, and aligned with healthcare budget requirements.

Theme 2: Preference for fairness, predictability, and transparency. All participants ($n = 17/17$, 100%) emphasized the need for transparent and predictable cost structures. In procurement contexts, multi-year budgets are common, and sudden cost fluctuations can complicate planning. Fairness was most often framed as proportionality between cost, value, and usage. Decision makers in all roles reported a willingness to accept higher prices if they are justified, stable, and supported by clear performance evidence. This emphasis on calculability and planning security was reflected in the comment of a senior microbiologist:

“Calculability through fixed price or licensed per result is scalable, has proven track record and good laboratory integration suitable for everyday use, including research.” (Senior Microbiologist)

Theme 3: Hybrid models as the preferred solution. Twelve participants ($n = 12/17$, 71%) advocated for hybrid models combining a fixed base fee with a variable component tied to clinically relevant metrics (e.g., per patient, per diagnostic episode). The base fee was viewed as critical to ensure service continuity, while the variable element enabled proportional cost adjustments as usage changes. Several participants highlighted that such structures allow for market differentiation without altering the underlying economic model, making them adaptable to both small practices and large hospital networks.

Domain 2: Reimbursement and Value Evidence

Themes 4 and 5 relate to expectations to demonstrate clinical value and the potential – but currently limited – role of outcome-based elements in pricing.

Theme 4: Outcome-based models—attractive but operationally challenging. Six participants ($n = 6/17$, 35%) found outcome-based pricing, linking payment to measurable improvements in diagnostic accuracy, efficiency or patient outcomes, ethically compelling and consistent with value-based care principles. However, methodological and legal obstacles were cited as major barriers. Outcomes are influenced by multiple concurrent factors, making it complex to attribute changes to a single AI system. Participants agreed that while outcome-linked reimbursement remains a strategic goal, significant evidence generation and regulatory adaptation would be required. This perspective is consistent with recent policy analyses that recommend linking reimbursement to demonstrated value, while cautioning that standardized assessment and harmonized criteria are prerequisites for a wider implementation [17]. Given these challenges, many participants argued for incremental innovation, not abandoning outcome-based ideas, but embedding AI into reimbursement pathways that are already operational.

Theme 5: Alignment with existing reimbursement frameworks (GOÄ/EBM). Ten participants ($n = 10/17$, 59%) stressed that aligning AI-enabled diagnostics with established frameworks such as GOÄ (Gebührenordnung für Ärzte; German physicians' fee schedule for privately billed services) and EBM (Einheitlicher Bewertungsmaßstab; the standard outpatient billing catalog for Germany's statutory health insurance) would remove major adoption barriers. This would allow AI services to be billed, processed, and audited within existing administrative workflows, thereby accelerating integration. Without integration into existing reimbursement frameworks, even technically mature AI systems can remain marginal in clinical practice [18].

Domain 3: Implementation Fit and Market Segmentation

Themes 6–8 highlight integration quality, workflow fit, market heterogeneity, and procurement practicality as determinants of willingness to adopt and pay for AI-enabled diagnostics.

Theme 6: Target market characteristics as determinants of pricing acceptance. Fourteen participants ($n = 14/17$, 82%) reported that price acceptance varies substantially with institutional size, resource availability, and procurement culture. Large hospital networks often prefer enterprise licensing to ensure budget stability, whereas smaller providers prefer modular offerings with low entry costs. From a market-economic perspective, this segmentation illustrates that a one-size-fits-all approach risks excluding key adopter groups.

Theme 7: Technical and organizational integration as a driver of willingness to pay. Fifteen participants ($n = 15/17$, 88%) identified seamless integration with existing IT infrastructure as a major determinant of perceived value. Systems that integrate without disrupting existing workflows were consistently rated more attractive and thus worth a higher cost. Participants frequently noted that the most successful integrations are “invisible” in daily operations, a sign of full compatibility

with existing processes. Participants also emphasized that integration can justify a higher price point, as a senior physician noted:

“Automatic forwarding of findings, integration into existing systems without manual intermediate steps and less visible applications increase acceptance - depth of integration justifies higher price.” (Senior Physician)

Theme 8: Bundling and cross-subsidization of modules. Eight participants ($n = 8/17$, 47%) described bundling strategies in which high-demand functionalities (i.e., modules with broad applicability and frequent use) are packaged together with niche capabilities (i.e., low-volume or specialty features) under a single commercial offering. The participants emphasized two main rationales. First, bundling reduces procurement and contracting complexity by replacing a long list of separately priced modules with a small number of comprehensible packages, which is easier to evaluate, purchase, and govern. Second, bundling enables cross-subsidization: revenues from widely used functions help cover the fixed costs of maintaining and validating niche features that are clinically important but would be difficult to sustain if priced solely on their low usage volume. In this sense, participants did not describe bundling as “hiding” costs, but as a pragmatic mechanism to ensure platform completeness, reduce transaction costs in procurement, and make adoption decisions easier for customers who otherwise must choose between an unwieldy menu of options.

Domain 4: Lifecycle Services and Equity

Themes 9 and 10 address service, training, long-term support, and equity-related concerns, reflecting the broader ethical and social implications of pricing strategies.

Theme 9: Service and support as integral pricing components. Sixteen participants ($n = 16/17$, 94%) agreed that pricing should explicitly include integration, training, and ongoing support. Although this may raise initial costs, participants argued that these services reduce operational risk and lower total cost of ownership. The expectation that service and support should be explicitly included in pricing was articulated, for example, by a resident in general medicine:

“For services including support there is a desire for an ‘all-inclusive package’ with clear responsibility assigned to one provider.” (Resident in general medicine)

Theme 10: Ethical, access-related, and societal considerations. Five participants ($n = 5/17$, 29%) highlighted the broader social implications of pricing strategies for diagnostic AI, particularly regarding equitable access between different types of healthcare provider. Several interviewees pointed out that smaller or rural laboratories often operate with considerably tighter budgets than large urban institutions, making them more vulnerable to price structures that scale with usage or require substantial upfront investment. Others emphasized that financial capacity to adopt AI varies markedly between institutions, and some providers are able to absorb more flexible pricing models more easily than others. From this perspective, participants warned that without explicit safeguards, AI adoption could widen existing regional disparities, as well-resourced university hospitals and large private laboratory groups might advance more rapidly than smaller regional clinics. Concerns about equity and

vulnerability of smaller providers were captured by a specialist in gastroenterology, who stated:

“Uncertainty about the future: Clinics/practices shy away from long-term commitments due to uncertainty regarding sources of financing. This is particularly difficult for individual practices and small non-profit organizations without a fixed investment budget.” (Specialist in gastroenterology)

Synthesis

Although the analysis is primarily qualitative, the thematic prevalence data show that the consensus was highest for transparency and predictability (Theme 2), integration-centric value perceptions (Theme 7), and explicit inclusion of service and support (Theme 9), each endorsed by more than 80% of participants. In contrast, outcome-linked models (Theme 4) and explicit equity safeguards (Theme 10) were less frequently prioritized, suggesting implementation gaps despite their normative appeal. Table 1 summarizes the identified themes and their relative agreement among the participants.

Table 1 Overview of identified themes and corresponding participant agreement

ID	Theme	in agreement (n)	Proportion (%)
1	Resistance to pure token-/pay-per-use models	13	76%
2	Preference for predictability and transparency	17	100%
3	Hybrid models as preferred solution	12	71%
4	Outcome-based models attractive but challenging	6	35%
5	Alignment with GOÄ/EBM reimbursement	10	59%
6	Target market characteristics shape acceptance	14	82%
7	Integration as driver of willingness to pay	15	88%
8	Bundling and cross-subsidization	8	47%
9	Service and support as integral components	16	94%
10	Ethical/access-related considerations	5	29%

Viewed through the domain structure introduced above, these prevalence patterns become easier to interpret and connect to design implications. Strong agreement clustered in *Pricing Architecture and Budget Predictability* (Domain 1) and *Implementation Fit and Market Segmentation* (Domain 3), reflecting participants’ emphasis on clinically legible units, predictable hybrid structures, and the primacy of integration and procurement practicality. By contrast, *Reimbursement and Value Evidence* (Domain 2) and *Lifecycle Services and Equity* (Domain 4) contain themes that were frequently acknowledged as important, but described as more difficult to operationalize uniformly across organizations without additional evidence, standardized assessment, or policy support. In line with the exploratory nature of the study, our interpretation prioritizes the diversity and nuance of stakeholder reasoning over numerical tallies. The counts complement the narrative by showing how widely a

theme was voiced, but they cannot capture silent agreement or disagreement and should therefore be read as descriptive context rather than quantitative evidence of salience.

Taken together, these findings indicate that the economic viability of AI-enabled DDSS depends on aligning pricing with the clinical and operational realities of health-care delivery: translating vendor-centric technical cost drivers into clinically legible units, ensuring procurement predictability through hybrid pricing, explicitly recognizing integration and lifecycle services, and embedding AI services into reimbursement pathways where feasible. Outcome-based and equity-driven elements remain strategically important, but require greater measurement maturity, governance clarity, and institutional readiness to be implemented at scale.

Discussion

The Discussion synthesizes the ten themes identified in the Results into stakeholder-informed design principles for pricing diagnostic AI and situates them within broader economic and policy debates on AI in healthcare. To address the central research question and improve readability while adhering to the journal's requirement of an unstructured Discussion, we first state the design principles explicitly and then elaborate on each through signposted paragraphs that connect empirical themes to practical and policy implications.

Across domains and themes, we derive six design principles for pricing AI-enabled DDSS. These principles correspond to the “actionable steps” previously summarized; here, we state them explicitly as the manuscript's design principles (DP1-DP6) and use them to guide the subsequent discussion:

DP1 : Use clinically legible billable units. Anchor pricing in units aligned with healthcare budgeting and clinical workload (e.g., per patient, per test, per episode) rather than technical meters.

DP2 : Prefer hybrid base & variable architectures to balance predictability and scalability. Combine a fixed base fee with a clinically defined variable component.

DP3 : Align pricing with reimbursement and audit logic wherever feasible. Enable coding, billing, and review within existing frameworks (e.g., GOÄ/EBM) and comparable systems internationally.

DP4 : Treat integration, training, and ongoing support as priced core components. Make lifecycle services explicit in commercial and reimbursement structures.

DP5 : Enable outcome-linked components when measurement readiness is sufficient, rather than making them a prerequisite for adoption. Where outcome linkage is pursued, couple it to robust evidence generation and standardized assessment.

DP6 : Incorporate equity-aware safeguards to avoid widening access gaps. Use tiering, staged rollout options, or co-funding mechanisms to support adoption among resource-constrained providers.

First, DP1 and DP2 (clinically legible units and hybrid architectures) are directly motivated by Themes 1–3. Participants consistently emphasized that token- or API-based pricing may be meaningful for internal vendor accounting but is not a clinically natural budgeting unit in healthcare organizations. Purely technical meters were described as opaque, difficult to forecast, and poorly aligned with reimbursement logic. Decision makers instead favored units that map to patient-centered financing—per case, per report, per test, or per episode—because these can be budgeted, benchmarked, and governed within established institutional processes. This finding aligns with prior arguments that procurement acceptance is shaped by transparency and predictable total cost of ownership, not solely by algorithmic performance [5, 6]. Hybrid architectures were framed as the pragmatic compromise that preserves planning security while accommodating scaling: a fixed base fee sustains readiness and continuity (e.g., monitoring, updates), while a clinically defined variable component preserves proportionality as use increases [14, 19]. In practical terms, DP1–DP2 imply a translation layer in which technically derived cost drivers remain largely internal and are mapped into clinically legible billing categories for purchasers.

Second, DP3 (reimbursement alignment and auditability) follows primarily from Themes 4–5. Even when organizations can finance pilots through innovation budgets, scaling typically requires stable billing and administrative pathways. Participants described alignment with established schedules such as GOÄ/EBM as removing friction in procurement, billing, and compliance review and as providing institutional legitimacy for AI-enabled diagnostic services. This aligns with conceptual work proposing AI-specific reimbursement logic within existing fee schedules and with concerns that generalist, multi-task AI challenges “one code per task” reimbursement models [18]. DP3 therefore emphasizes that pricing should be designed for real-world coding, billing, and auditability rather than merely justified as a technical subscription.

Third, DP4 (integration and lifecycle services as core priced components) is supported by Themes 6–9, which indicate that stakeholders value AI systems as implementation-and-service bundles rather than as software licenses alone. Integration quality (Theme 7) was repeatedly described as the difference between theoretical benefit and routine usability, and participants emphasized that poorly integrated systems impose hidden costs through workflow disruptions and additional staff effort. This is consistent with implementation research emphasizing that real-world value depends on organizational fit, maintenance, and lifecycle effort [8, 14]. DP4 therefore recommends making integration, training, customization, and ongoing support explicit, priced components of the offering. At the same time, Theme 7 reveals a tension between the preference for “invisible” integration and service-logic perspectives on value creation. In Grönroos’ service logic, value is co-created in use when organizations adapt processes and actively work with a service [29]. Our findings suggest that some value potential may remain unrealized if institutions aim exclusively to avoid workflow change, reinforcing the need to resource change management and training explicitly as part of the service package.

Fourth, DP5 (outcome linkage as an evidence-enabled option) reflects Theme 4: participants found outcome-based pricing ethically compelling and aligned

with value-based care, but operationally difficult due to attribution and measurement constraints. Outcomes often reflect multiple concurrent interventions, and organizations may lack harmonized data pipelines or accepted evaluation standards. This is consistent with policy analyses that support outcome-linked reimbursement in principle but emphasize the need for standardized assessment and evidence generation [17]. DP5 therefore frames outcome linkage as a feasible add-on when measurement readiness exists, not as a prerequisite that would block early adoption. Operationally, this implies a staged pathway: adopt clinically legible hybrid pricing (DP1–DP2), embed lifecycle services (DP4), and then add outcome-linked components in contexts where impact can be measured credibly and governance frameworks permit.

Fifth, DP6 (equity-aware safeguards) is motivated by Theme 10 and is reinforced by broader health-equity debates. Participants cautioned that pricing structures can widen disparities if they implicitly favor large, well-resourced organizations; smaller and rural providers often have tighter budgets and higher sensitivity to financial uncertainty. DP6 therefore supports tiered pricing, staged adoption packages, and co-funding or shared infrastructure mechanisms to reduce adoption barriers for resource-constrained providers. This aligns with the broader concern that digital innovations can diffuse unevenly and exacerbate inequities without deliberate policy design [22].

Beyond these near-term principles, participants' emphasis on ongoing support (Theme 9) intersects with emerging regulatory obligations, including monitoring and post-market activities under AI governance frameworks such as the European Union AI Act [27]. As per-inference costs decline due to scale effects and optimization, pricing and reimbursement will increasingly depend on how fixed lifecycle obligations and potential surplus distribution are governed across payers, providers, and vendors [13, 19]. In parallel, there is a trajectory toward more autonomous and multi-task systems, which may require movement from discrete per-decision fees to bundled or episode-based models and the development of AI-specific modifiers for autonomy and scope of contribution [18]. These forward-looking dynamics strengthen the rationale for DP3 (auditability and coding logic) and DP5 (measurement readiness for outcome linkage), while increasing the importance of DP6 to prevent uneven capture of efficiency gains.

Several limitations should be considered when interpreting these findings. The study synthesizes perspectives from $n = 17$ stakeholders in clinical and administrative roles in German-speaking countries; while diverse in roles and institutions, the sample is not statistically representative and transferability to other health systems must be contextualized. The interviews addressed diagnostic AI in general rather than a single technology, specialty, or deployment scenario; pricing preferences may vary across contexts (e.g., radiology, laboratory triage, administrative decision support) with different risk profiles and value pathways. The gender distribution was markedly imbalanced (16 male, 1 female), which may under-represent perspectives of women decision makers. Finally, the qualitative design prioritizes depth over quantification; we did not elicit willingness-to-pay parameters or test pricing structures in prospective procurement settings. Future research should therefore (i) evaluate

hybrid architectures in procurement pilots, (ii) develop attribution-ready measurement frameworks for outcome-linked components, and (iii) assess equity safeguards across small-provider and rural settings to avoid access gaps, ideally complemented by formal HTA studies to quantify integration effort, workflow improvements, and outcome-linked value.

Motivated by the gap outlined in the *Introduction*, this article examined how healthcare decision makers think pricing for AI-enabled diagnostic decision support should be structured so that it fits procurement and reimbursement processes while accommodating AI-specific technical and organizational cost drivers and usage patterns. The six design principles (DP1–DP6) provide a stakeholder-informed blueprint for clinically meaningful, predictable, and technically faithful AI pricing: anchor billable units in clinical categories, use hybrid base+variable architectures, align with reimbursement and audit logic, explicitly price integration and lifecycle services, pursue outcome linkage only where measurement readiness exists, and incorporate equity-aware safeguards. These principles translate into actionable implications for policymakers, payers, and vendors. In particular, embedding AI services into established reimbursement structures (e.g., GOÄ/EBM) with AI-specific modifiers can enhance transparency and facilitate processing and auditability; hybrid pricing can balance predictability and scalability without reverting to opaque technical meters; and explicit recognition of integration and support can reduce implementation risk and improve total cost of ownership. As AI capabilities evolve, deliberate policy will be needed to ensure that efficiency gains are distributed fairly and that adoption strengthens rather than erodes equitable access to diagnostic innovation.

Methods

Healthcare decision makers shape adoption of diagnostic AI by determining procurement, integration priorities, and long-term financing. We therefore used qualitative interviews to elicit how such decision makers evaluate pricing structures and reimbursement fit for AI-enabled diagnostic decision support systems. This study adopts a qualitative approach to identify practical considerations, unmet needs, and expectations for future pricing strategies from the perspective of decision makers in healthcare.

Study design and motivation

We employed an exploratory qualitative research design based on semi-structured expert interviews to examine how healthcare decision makers assess and accept pricing models for AI-enabled diagnostic decision support. A qualitative design is particularly suited for this topic because the evidence base is still emerging and because pricing acceptance depends on institutional constraints, procurement practices, and experiential reasoning that are difficult to capture through standardized instruments alone. Semi-structured interviews allow comparability across participants while enabling probing of context-specific considerations and trade-offs.

Given persistent misalignment between AI pricing approaches and healthcare budgeting and administrative logic [5, 6], we used expert interviews to elicit how decision makers translate economic and technical concepts into procurement decisions, which design features they perceive as feasible, and which conditions they view as prerequisites for sustainable diffusion.

All interviews were conducted digitally via a secure video conferencing platform. The audio material was recorded and subsequently transcribed in full, preserving verbatim all spoken content. To balance depth with participant availability, each interview lasted between 45 and 60 minutes. In cases where participants requested additional time, follow-up sessions were offered.

This study draws on expertise in clinical decision support, health economics, and integration of healthcare IT. To mitigate potential bias from previous industry and academic exposure to DDSS, the interview guide was piloted with external reviewers, and coding and theme synthesis were performed after data collection and iteratively checked against the transcripts. The interpretations were iteratively refined to ensure that the analytical choices reflected the perspectives of the participants rather than preconceived pricing preferences.

Participant selection and sampling strategy

A purposive sampling strategy was applied to ensure broad coverage of professional roles, care settings, and varying degrees of familiarity with Diagnostic Decision Support Systems (DDSS). The main inclusion criterion was substantial professional involvement in decision-making processes related to diagnostics, healthcare technology, or digital health integration. This included both direct clinical roles, such as physicians and specialists actively engaged in patient care, and adjacent positions with strategic influence, for example, in health IT management, biomedical engineering, consulting, or reimbursement policy.

Participants were identified through three complementary recruitment channels: (i) established professional networks in laboratory medicine, clinical informatics, and health technology, (ii) academic collaborations and contacts established through conferences and expert panels, and (iii) targeted outreach to healthcare institutions known to be exploring or implementing AI-based diagnostic tools. This multipronged approach ensured access to both individuals with direct operational experience and those involved in high-level strategic or policy discussions.

The final sample comprised $n = 17$ experts from Germany, Austria, and Switzerland, aged 26 to 77 years (median: 37 years). The sample is demographically skewed (16 male, 1 female) and was skewed toward senior/leadership positions, which may have influenced participants' perceptions of feasibility of procurement and reimbursement alignment. Eleven participants had a clinical background, representing a broad range of specialties, including laboratory medicine, internal medicine, gastroenterology, radiation therapy, general practice, transfusion medicine, and microbiology. Several of these clinicians occupied senior leadership positions, such as department heads, senior consultants, or medical directors, combining direct patient care with administrative, organizational, and strategic responsibilities. The six non-clinical participants brought expertise in health IT management, biomedical engineering, digital

health consulting, life sciences business development, and reimbursement advisory. Their roles frequently involved close collaboration with clinical teams, particularly in the context of system procurement, integration projects, and strategic planning for the adoption of AI-driven healthcare solutions. Institutional affiliations spanned large public university hospitals, private diagnostic laboratory chains, regional group practices, and health technology companies. Specifically, the sample included 5 participants from large public hospitals, 3 from private diagnostic laboratory chains, 3 from regional group practices, and 6 from health technology companies. This diversity of organizational contexts ensured the inclusion of perspectives shaped by different operational scales, resource availability, and regulatory environments.

The main attributes of the sample are included in an anonymized interview summary table (Supplementary Data 1), which is also publicly accessible on Zenodo (<https://doi.org/10.5281/zenodo.17116813>).

To avoid conflicts of interest in data collection, none of the interviewees were employees of medicalvalues GmbH or direct colleagues in the same immediate teams as the academic authors at their universities. Two of the authors (Jan Kirchhoff and Fabian Berns) are employed by a diagnostic AI company (medicalvalues GmbH); to mitigate potential bias, we used a neutral interview protocol, emphasized the academic nature of the study to all participants, and conducted targeted spot checks by a second author during analysis.

The experience with DDSS among participants varied widely. Some reported daily operational use, often combined with active participation in in-house system development or optimization. Others described indirect exposure through procurement processes, pilot projects, or system integration initiatives. A smaller subset indicated mainly conceptual familiarity, such as evaluation of DDSS from a policy, strategy, or market perspective, without regular application in daily workflows. This diversity in experience was intentionally sought to capture both the practical realities of experienced adopters and the perspectives of potential future users.

We continued recruitment until thematic saturation was reached, which meant that additional interviews were no longer generating new conceptual insights; by approximately the fifteenth interview, no substantially new themes were emerging, and the final two interviews confirmed saturation.

All invited participants received essential study information in advance and participation was voluntary. Financial or material incentives were not provided.

Interview structure and content

The interview guide was developed based on an initial review of the literature on AI pricing models in healthcare and refined through discussions with domain experts. We used a semi-structured format with a fixed set of core questions that were asked of all participants, complemented by optional probes. At the beginning of each interview, participants were invited (only as much as they felt comfortable) to briefly describe their age, gender, and professional background. The main part of the guide was organized into four thematic blocks:

1. **Context and experience:** Understanding of AI-based diagnostic decision support systems (DDSS) and prior exposure to such tools.
2. **Perceived costs and benefits:** Expectations about implementation effort, clinical and economic value, and suitable indicators to assess impact.
3. **Pricing models and reimbursement logic:** Views on different pricing structures (e.g., flat licensing, usage-based, complexity-based) and their compatibility with budgeting and regulatory requirements.
4. **Future requirements:** Key requirements and recommendations for future pricing models for DDSS.

Within each block, open-ended core questions (for example, asking participants which pricing models they considered most appropriate in their setting and why) were asked in a consistent order across interviews to support comparability. Follow-up prompts were used flexibly to clarify responses or generate concrete examples, allowing important topics to be covered systematically while preserving the depth and nuance typical of qualitative interviewing.

Transcription, anonymization, and data analysis

Participants were informed about the measures taken to protect their privacy. All interviews were manually transcribed by the primary author on local encrypted devices. Domain-specific terminology was standardized during transcription, and non-lexical fillers were removed only when meaning was not affected. In accordance with the General Data Protection Regulation (GDPR) and ethical safeguards, speakers were replaced by role-based identifiers and identifiable entities were generalized prior to analysis.

In determining sample adequacy, we drew on the concept of “information power” [24], which posits that studies with a focused objective and highly experienced participants require fewer interviews to generate robust analytical insights. Given the seniority, domain expertise, and diversity of roles represented in our sample, the information power was high, supporting our conclusion that 17 interviews were sufficient.

The analysis followed a mixed deductive–inductive thematic analysis [25], with the aim of obtaining stakeholder-informed insights. Our approach was based on the six-phase procedure described by Braun and Clarke [25], including familiarization with the data, generation of initial codes, construction and review of candidate themes, definition and naming of themes, and final synthesis. The statements were first organized according to the predefined blocks of the interview guide (context, perceived costs and benefits, evaluation of pricing models, and future requirements). The annotation combined a priori categories derived from the interview guide with inductively generated codes that captured emergent issues such as reimbursement alignment, integration effort, and equity considerations. A compact codebook (labels, definitions, inclusion, and exclusion criteria) was calibrated on a pilot transcript independently coded by two authors; disagreements were discussed to align interpretations, after which the refined codebook was applied by the primary author with targeted spot checks by a second author.

In line with the Framework Method of Gale [26], we then constructed a participant-by-code matrix to support constant comparison between cases and the aggregation of related material. The draft theme statements were iteratively refined against anonymized excerpts until they provided a coherent and concise representation of the underlying data. To enhance transparency and trustworthiness, we also selected a small number of anonymized illustrative quotations from the interview materials and report them in the Results to exemplify key themes; these quotations are consistent with the summarized responses documented in the anonymized interview summary table (see Data availability).

In the report, prevalence counts (e.g., $n = 13/17$) indicate how many different participants explicitly endorsed a theme at least once. These tallies are descriptive aids grounded in the coded matrix and should not be interpreted as inferential statistics or as indicators of thematic importance. Given the semi-structured, open-ended format, the absence of a theme in an individual interview simply reflects that it was not explicitly mentioned in that conversation.

Ethics and consent to participate

The study protocol was reviewed by the responsible institutional contact at Neu-Ulm University of Applied Sciences. In accordance with the guidelines of the joint Ethics Committee of the Bavarian Universities, expert interviews involving adults, without patients or vulnerable populations, without collection of sensitive personal data and relying solely on fully anonymized analysis do not require formal ethics committee review. As the present study meets these criteria, no formal ethics approval and no reference number were required. The study adhered to the principles of the Declaration of Helsinki. All participants received a written information sheet detailing the study aims, procedures, data protection measures, and their rights. Written informed consent was obtained prior to each interview, including consent for audio recording and anonymized use of their statements for research and publication. Participation was voluntary and participants could withdraw at any time without consequences.

Declarations

Data Availability. The anonymized interview summary table of the stakeholder responses that support the findings of this article is provided as Supplementary Data 1 (Excel file) and is also publicly available on Zenodo: <https://doi.org/10.5281/zenodo.17116813>. Full interview transcripts are not publicly available due to privacy and confidentiality restrictions and in accordance with GDPR. Additional materials may be made available from the corresponding author upon reasonable request.

Code Availability. No custom code was used to generate or analyze the data in this study; therefore, the availability of the code is not applicable.

Acknowledgements. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions. J.K. conceived the study, served as lead author, and coordinated the design and execution of the stakeholder interviews. J.S. supervised the overall study design and organizational framework. F.B. contributed to the preparation of interview materials, data tabulation, and review of interview transcripts. C.S. provided methodological guidance and critical feedback on the scientific approach. All authors contributed to the interpretation of results and reviewed the final manuscript. The authors declare that the AI-based software 'Writefull' was used to improve the language and grammar of the manuscript. The tool did not generate or alter content, data analysis, or interpretation.

Competing Interests. JK and FB are employed by medicalvalues GmbH (Karlsruhe, Germany). The company had no role in the design, conduct, analysis, or decision to publish this work. All authors declare no additional financial or non-financial competing interests.

References

- [1] Tu, T., *et al.* Towards conversational diagnostic artificial intelligence. *Nature* **642**, 442–450 (2025). doi: 10.1038/s41586-025-08866-7.
- [2] Sutton, R. T., *et al.* An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digital Medicine* **3**, 17 (2020). doi: 10.1038/s41746-020-0221-y.
- [3] Hazarika, I. Artificial intelligence: opportunities and implications for the health workforce. *International Health* **12**, 241–245 (2020). doi: 10.1093/inthealth/ihaa007.
- [4] Alowais, S. A., *et al.* Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education* **23**, 689 (2023). doi: 10.1186/s12909-023-04698-z.
- [5] Chen, M. M. & Nicola, G. N. Who Will Pay for AI? *Radiology: Artificial Intelligence* **3**, e210030 (2021). doi: 10.1148/ryai.2021210030.
- [6] Venkatesh, K. P., *et al.* Leveraging reimbursement strategies to guide value-based adoption and utilization of medical AI. *npj Digital Medicine* **5**, 112 (2022). doi: 10.1038/s41746-022-00662-1.
- [7] Elhaddad, M. & Hamam, S. AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential. *Cureus* **16**, e57728 (2024). doi: 10.7759/cureus.57728.
- [8] Abell, B., *et al.* Identifying barriers and facilitators to successful implementation of computerized clinical decision support systems in hospitals: a NASSS framework-informed scoping review. *Implementation Science* **18**, 32 (2023). doi: 10.1186/s13012-023-01287-y.

- [9] Hiltawsky, K. & Boll, S., *et al.* KI-Geschäftsmodelle für die Gesundheit: Innovation stärken, Finanzierung gestalten. *Plattform Lernende Systeme* (2022). doi: 10.48669/pls.2022-3.
- [10] Borowiec, D., *et al.* Trimmer: Cost-Efficient Deep Learning Auto-tuning for Cloud Datacenters. In: *2022 IEEE 15th International Conference on Cloud Computing (CLOUD)*. 374–384 (IEEE, 2022). doi: 10.1109/CLOUD55607.2022.00061.
- [11] Guo, X., Yang, Z. & Yi, J. Design and implementation of efficient distributed deep learning model inference architecture on serverless computation. *Applied and Computational Engineering* **68**, 347–355 (2024). doi: 10.54254/2755-2721/68/20241535.
- [12] Kumar, N. *Pricing AI Products: The Monetization Playbook for Startup Founders*. (Nitin Kumar, 2025). ISBN: 9798992302400.
- [13] Hirosawa, T., *et al.* Evaluating ChatGPT-4's Diagnostic Accuracy: Impact of Visual Data Integration. *JMIR Medical Informatics* **12**, e55627 (2024). doi: 10.2196/55627.
- [14] Leusder, M., *et al.* Cost measurement in value-based healthcare: a systematic review. *BMJ Open* **12**, e066568 (2022). doi: 10.1136/bmjopen-2022-066568.
- [15] Bratan, T., *et al.* Unterstützung ärztlicher und pflegerischer Tätigkeit durch KI: Handlungsempfehlungen für eine verantwortbare Gestaltung und Nutzung. *Bundesgesundheitsblatt* **67**, 1039–1046 (2024). doi: 10.1007/s00103-024-03918-1.
- [16] Giansanti, D. & Pirrera, A. Integrating AI and Assistive Technologies in Healthcare: Insights from a Narrative Review of Reviews. *Healthcare* **13**, 556 (2025). doi: 10.3390/healthcare13050556.
- [17] Tarricone, R., Petracca, F. & Weller, H.-M. Towards harmonizing assessment and reimbursement of digital medical devices in the EU through mutual learning. *npj Digital Medicine* **7**, 268 (2024). doi: 10.1038/s41746-024-01263-w.
- [18] Dogra, S., Silva, E. & Rajpurkar, P. Reimbursement in the age of generalist radiology artificial intelligence. *npj Digital Medicine* **7**, 350 (2024). doi: 10.1038/s41746-024-01352-w.
- [19] Steinbeck, V., Freseemann, M. L. & Busse, R. “Value-based healthcare” als Möglichkeit zur PatientInnensteuerung auf individueller und Systemebene. *Die Innere Medizin* **65**, 890–898 (2024). doi: 10.1007/s00108-024-01767-3.
- [20] Khanna, N. N., *et al.* Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare* **10**, 2493 (2022). doi: 10.3390/healthcare10122493.

- [21] Huang, K. A., Choudhary, H. K. & Kuo, P. C. Artificial Intelligent Agent Architecture and Clinical Decision-Making in the Healthcare Sector. *Cureus* **16**, e64115 (2024). doi: 10.7759/cureus.64115.
- [22] Abràmoff, M. D., *et al.* Considerations for addressing bias in artificial intelligence for health equity. *npj Digital Medicine* **6**, 170 (2023). doi: 10.1038/s41746-023-00913-9.
- [23] Schmidt, L., *et al.* The three-year evolution of Germany's Digital Therapeutics reimbursement program and its path forward. *npj Digital Medicine* **7**, 139 (2024). doi: 10.1038/s41746-024-01137-1.
- [24] Malterud, K., Siersma, V. D. & Guassora, A. D. Sample size in qualitative interview studies: guided by information power. *Qualitative Health Research* **26**, 1753–1760 (2016). doi: 10.1177/1049732315617444.
- [25] Braun, V. & Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* **3**, 77–101 (2006). doi: 10.1191/1478088706qp063oa.
- [26] Gale, N. K., Heath, G., Cameron, E., Rashid, S. & Redwood, S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology* **13**, 117 (2013). doi: 10.1186/1471-2288-13-117.
- [27] European Parliament & Council. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union* (OJ L, 2024/1689, 12.7.2024). <http://data.europa.eu/eli/reg/2024/1689/oj>.
- [28] Kirchhoff, J., Schieder, C., Berns, F. & Schobel, J. Valuing diagnostic AI: a structured reimbursement model for learning healthcare systems. *Frontiers in Digital Health* **7**, 1642750 (2025). doi: 10.3389/fdgth.2025.1642750.
- [29] Grönroos, C. Value co-creation in service logic: A critical analysis. *Marketing Theory* **11**, 279–301 (2011). doi: 10.1177/1470593111408177.